

PCT

世界知的所有権機関
国際事務局
特許協力条約に基づいて公開された国際出願



<p>(51) 国際特許分類6 C07K 14/47, C12N 15/12, 5/10, C12P 21/02, C07K 16/18, C12P 21/08</p>	<p>A1</p>	<p>(11) 国際公開番号 WO99/57143</p> <p>(43) 国際公開日 1999年11月11日(11.11.99)</p>
<p>(21) 国際出願番号 PCT/JP99/02340</p> <p>(22) 国際出願日 1999年4月30日(30.04.99)</p> <p>(30) 優先権データ 特願平10/137631 1998年4月30日(30.04.98) JP</p> <p>(71) 出願人 (米国を除くすべての指定国について) 株式会社 中外分子医学研究所(CHUGAI RESEARCH INSTITUTE FOR MOLECULAR MEDICINE, INC.)[JP/JP] 〒300-4101 茨城県新治郡新治村永井153番地2 Ibaraki, (JP)</p> <p>(72) 発明者 ; および (75) 発明者 / 出願人 (米国についてのみ) ジョーンズ マイケル エイチ(JONES, Michael H.)[GB/JP] 〒300-4101 茨城県新治郡新治村永井153番地2 株式会社 中外分子医学研究所内 Ibaraki, (JP)</p> <p>(74) 代理人 清水初志, 外(SHIMIZU, Hatsushi et al.) 〒300-0847 茨城県土浦市御町1-1-1 関鉄つくばビル6階 Ibaraki, (JP)</p>		<p>(81) 指定国 AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, 欧州特許 (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI特許 (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG), ARIPO特許 (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), ユーラシア特許 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM)</p> <p>添付公開書類 国際調査報告書</p>
<p>(54) Title: TRANSCRIPTIONAL REGULATORY FACTOR</p> <p>(54) 発明の名称 転写調節因子</p> <p>(57) Abstract BLAST detection is performed on EST data base by using various base sequences encoding known bromo-domain motives and thus several ESTs likely encoding bromo-domain gene are found out. Next, testicular cDNAs are PCR-cloned by using primers designed based on the sequence of EST(W17142) which are one of the ESTs found out above. By using the thus obtained PCR product as a probe, the testicular cDNA library is screened. Further, the testicular cDNA library is screened again by using the obtained cDNA clone as a probe, thereby successfully isolating a full-length cDNA corresponding to EST(W17142). The protein encoded by the thus isolated cDNA has several regions and domains conserved in the transcriptional regulatory factor, in addition to the bromo-domain. Moreover, the protein undergoes interactions with a protein seemingly relating to the chromatin-mediated transcriptional regulatory mechanism and transcriptional coactivator.</p>		

(57)要約

既知のプロモドメインモチーフをコードする種々の塩基配列を用い、EST データベースに対して BLAST 検索を行った結果、プロモドメイン遺伝子をコードする可能性のある幾つかの EST を見出した。次いで、これら EST の一つである EST(W17142)の配列を基にデザインしたプライマーを利用して精巢 cDNA の PCR クローニングを行い、さらに該 PCR 産物をプローブとした精巢 cDNA ライブラリーのスクリーニングおよび得られた cDNA クローンをプローブとした精巢 cDNA ライブラリーの再スクリーニングにより、ESTW17142 に対する全長 cDNA を単離することに成功した。単離した cDNA がコードするタンパク質は、がプロモドメイン以外に、転写調節因子に保存されたいくつかの領域やドメインを有していた。また、クロマチン介在転写制御機構への関与していることが示唆されるタンパク質や転写コアクチベーターと相互作用していた。

PCTに基づいて公開される国際出願のパンフレット第一頁に掲載されたPCT加盟国を同定するために使用されるコード(参考情報)

AE アラブ首長国連邦	DM ドミニカ	KZ カザフスタン	RU ロシア
AL アルバニア	EE エストニア	LC セントルシア	SD スーダン
AM アルメニア	ES スペイン	LI リヒテンシュタイン	SE スウェーデン
AT オーストリア	FI フィンランド	LK スリ・ランカ	SG シンガポール
AU オーストラリア	FR フランス	LR リベリア	SI スロヴェニア
AZ アゼルバイジャン	GA ガボン	LS レソト	SK スロヴァキア
BA ボスニア・ヘルツェゴビナ	GB 英国	LT リトアニア	SL シエラ・レオネ
BB バルバドス	GD グレナダ	LU ルクセンブルグ	SN セネガル
BE ベルギー	GE グルジア	LV ラトヴィア	SZ スワジランド
BF ブルキナ・ファソ	GH ガーナ	MA モロッコ	TD チャード
BG ブルガリア	GM ガンビア	MC モナコ	TG トーゴ
BJ ベナン	GN ギニア	MD モルドヴァ	TJ タジキスタン
BR ブラジル	GW ギニア・ビサウ	MG マダガスカル	TZ タンザニア
BY ベラルーシ	GR ギリシャ	MK マケドニア旧ユーゴスラヴィア	TM トルクメニスタン
CA カナダ	HR クロアチア	共和国	TR トルコ
CF 中央アフリカ	HU ハンガリー	マリ	TT トリニダード・トバゴ
CG コンゴ	ID インドネシア	MN モンゴル	UA ウクライナ
CH スイス	IE アイルランド	MR モーリタニア	UG ウガンダ
CI コートジボアール	IL イスラエル	MW マラウイ	US 米国
CM カメルーン	IN インド	MX メキシコ	UZ ウズベキスタン
CN 中国	IS アイスランド	NE ニジェール	VN ヴェトナム
CR コスタ・リカ	IT イタリア	NL オランダ	YU ユーゴスラビア
CU キューバ	JP 日本	NO ノールウェー	ZA 南アフリカ共和国
CY キプロス	KE ケニア	NZ ニュー・ジーランド	ZW ジンバブエ
CZ チェコ	KG キルギスタン	PL ポーランド	
DE ドイツ	KP 北朝鮮	PT ポルトガル	
DK デンマーク	KR 韓国	RO ルーマニア	

明細書

転写調節因子

技術分野

本発明は、プロモドメインを有する新規な転写調節因子およびその遺伝子に関する。

背景技術

プロモドメインは、転写調節因子に見られる特徴的なアミノ酸のモチーフであり、他の転写調節因子などとの相互作用に関与すると考えられている。プロモドメインを有するタンパク質は、通常、1 個または 2 個 (Tamkun JW et al. (1992). Nuc. Acids. Res., 20, 2603; Haynes SR et al. (1992). Nuc. Acids. Res., 20, 2603)、あるいは 5 個 (Nicolas RH and Goodwin GH. (1996). Gene, 175 (12), 233-240) のプロモドメインモチーフを含んでいる。このモチーフの見られる動物は広範囲にわたっており、例えば、ショウジョウバエ(drosophila)におけるホメオ遺伝子 (Digan ME et al. (1986). Dev. Biol., 114, 161-169; Tamkun JW et al. (1992). Cell, 68, 561-572) や酵母 (Winston F et al. (1987). Genetics, 115, 649-656; Laurent BC et al. (1991). Proc. Nat. Acad. Sci., USA 88, 2687-2691) および哺乳動物 (Denis GV, and Green MR. (1996). Genes and Devel., 10, 261-271; Yang X J et al. (1996). Nature, 382, 319-324) の転写調節遺伝子などで同定されている。最近の報告 (Jeanmougin et al. (1997). Trends Biochem. Sci. 22, 151-153) によれば、13 のヒトの遺伝子を含む 37 のプロモドメイン遺伝子がデータベースに登録されている。さらには 59-63 アミノ酸残基のプロモドメインモチーフに加え、モチーフに近接する配列も構造的に保存されており、それらはさらに長い 110 アミノ酸に 4 つの α -ヘリックス (Z,

A, B および C)がコードされていることが報告されている。

プロモドメインを有する転写調節因子を比較すると、これらのすべては活発に増殖している細胞でシグナル依存性の転写を調節している (Tamkun JW et al.(1992). Cell, 68, 561-572、Haynes SR et al. (1992). Nuc. Acids. Res., 20, 2603)。この特徴は、プロモドメインを有するタンパク質をコードする遺伝子の正常な制御が行われない場合に発癌する可能性があることを示唆している。現に6つのヒトプロモドメイン遺伝子との関連が実験的に証明されている。これら遺伝子の中の3つ HRX/ALL-1 (Tkachuk et al.(1992). Cell, 71, 691-700、Gu et al.(1992) Cell, 71, 701-708)、TIF1 (Miki et al.(1991) Proc. Nat. Acad. Sci., 88,5167-5171、Le Douarin et al.(1995) EMBO J. 14, 2020-2033) や CBP (Borrow et al. (1996). Nature Genet., 14, 33-41) は白血病の遺伝子切断点と関連している。また、これらの3つのタンパク質は全て C4HC3 (別名 PHD/LAP/TRX)ジンクフィンガー(Aasland et al.(1995). Trends Biochem. Sci., 20, 56-59、Koken et al.(1995). CR Acad. Sci. III, 318, 733-739、Saha et al.(1995). Proc. Nat. Acad. Sci.,92, 9737-9741)を保持している。さらに、CBP/P300 は p53 (Gu et al.(1997) Nature 387, 819-823、Lill et al.(1997) Nature 387,823-827)や他の様々な転写因子と相互作用しそれらを制御しているという知見もあり、このことは CBP やそれに非常に相同性がある P300 が癌においては鍵となる役割を担っていることを示唆するものである。

他の3つの遺伝子は様々なかたちで癌との関連が示唆されている。BRG1 は癌抑制遺伝子である網膜芽細胞腫タンパク Rb (Dunaief et al.(1994) Cell 79, 119-130)と相互作用し、単層化し、増殖が停止した細胞の形成を誘導することによって癌抑制活性を示す。RING3 は、ショウジョウバエ (*Drosophila*) 成長制御タンパク fsh (Haynes et al.(1989). Dev. Biol., 134, 246-257)と相同性があり、核内自己リン酸化活性を有するセリノースレオニンキナーゼである。この活性は慢性和急性リンパ性白血病での増殖期と関連があることが報告され

ている(Denis and Green,(1996). Genes and Devel., 10, 261-271)。P/CAF については、E1A と p300/CBP との相互作用を阻害することが知られている(Yang et al,(1996). Nature, 382, 319-324)。HeLa 細胞で P/CAF を外来性に発現させると、細胞周期の進行が阻害される。これはすなわち P/CAF が p300/CBP へ結合することにより、E1A の転写制御を乱してしまうというメカニズムに由来しているものと考えられる。また p300/CBP (Bannister and Kouzarides,(1996). Nature, 384, 641-643)と同様、P/CAF はヒストンアセチルトランスフェラーゼ活性を持つことも報告されている(Yang et al.(1996). Nature, 382, 319-324)。

このようにプロモドメインを有する転写調節因子は、その調節異常が種々の疾患、特に、癌やその他の増殖性疾患と密接に関与していると考えられる。このためプロモドメインを有する転写調節因子は、癌やその他の増殖性疾患の治療のための新しい標的として近年注目されている。

発明の開示

本発明はプロモドメインを有する新規な転写調節因子およびその遺伝子、並びにこれらの製造方法およびこれらを利用した医薬品候補化合物のスクリーニング方法を提供する。

本発明者らは上記課題を解決すべく、まず既知のプロモドメインモチーフをコードする種々の塩基配列を用い、EST データベースに対して BLAST 検索を行った。その結果、テトラヒメナ サーモフィラ (*Tetrahymena thermophila*) HAT A1 遺伝子の塩基配列を用いた検索により、プロモドメイン遺伝子をコードする可能性のあるいくつかの EST を見出した。これら EST の一つである胎児肺 cDNA ライブラリー由来の EST(W17142)は、未知の遺伝子をコードしていた。そこで、次に、ESTW17142 に対する全長 cDNA の単離を試みた。具体的には、ESTW17142 の配列を基にプライマーをデザインし、精巢の cDNA を鋳型にポリメラーゼ連鎖反応を行って増幅産物を得た。次いで、この増幅産物をプローブと

して精巢 cDNA ライブラリーのスクリーニングを行い、さらに上記 EST の配列を有する cDNA クローンを有したライブラリーの再スクリーニングを行った。その結果、ESTW17142 に対する全長 cDNA を単離することに成功した。また、本発明者らは、単離した cDNA がコードするタンパク質の構造につき解析を行い、該タンパク質がプロモドメイン以外に、転写調節因子に保存されたいくつかの領域やドメインを有することを見出した。

また、本発明者等は、単離した cDNA がコードするタンパク質が、クロマチン介在転写制御機構の一連の過程に関与していることが示唆される hSNF2H および hSNF2L や、種々の核受容体 (VDR、RAR) のリガンド結合ドメインやウィルス腫瘍タンパク質 (oncoprotein) である Ski と相互作用する転写コアクチベーター NCoA-62/Skip と相互作用することを見出した。

本発明者等により見出された転写調節因子およびその遺伝子は、該転写調節因子とこれに相互作用する因子との結合の阻害する化合物やその活性を制御する化合物のスクリーニングへの利用が可能であり、このスクリーニングにより単離される化合物は医薬品への応用が期待される。

即ち、本発明は、プロモドメインを有する新規な転写調節因子およびその遺伝子、並びにそれらの製造およびそれらを利用した関連因子および医薬品候補化合物のスクリーニングに関し、より具体的には、

1. 配列番号：1または10に記載のアミノ酸配列からなるタンパク質、
2. 配列番号：1または10に記載のアミノ酸配列において1若しくは複数のアミノ酸が置換、欠失、付加、および/または挿入されたアミノ酸配列からなり、プロモドメインを有する転写調節因子、
3. 配列番号：1または10に記載のアミノ酸配列において1若しくは複数のアミノ酸が置換、欠失、付加、および/または挿入されたアミノ酸配列からなり、hSNF2H、hSNF2L、NCoA-62/Skip およびそれらの他種相同体からなる群より選択されるタンパク質に結合する活性を有するタンパク質、

4. 配列番号：2または9に記載の塩基配列からなるDNAとハイブリダイズするDNAがコードし、プロモドメインを有する転写調節因子、
5. 配列番号：2または9に記載の塩基配列からなるDNAとハイブリダイズするDNAがコードし、hSNF2H、hSNF2L、NCoA-62/Skip およびそれらの他種相同体からなる群より選択されるタンパク質に結合する活性を有する転写調節因子、
6. (1)から(5)のいずれかに記載の転写調節因子をコードするDNA、
7. 配列番号：2または9に記載の塩基配列のコード領域を含む、(6)に記載のDNA、
8. (6)または(7)に記載のDNAを含むベクター、
9. (6)または(7)に記載のDNAを発現可能に保持する形質転換体、
10. (9)に記載の形質転換体を培養する工程を含む、(1)から(5)のいずれかに記載の転写調節因子の製造方法、
11. (1)から(5)のいずれかに記載の転写調節因子に結合する抗体、
12. (1)から(5)のいずれかに記載の転写調節因子に結合する活性を有する化合物のスクリーニング方法であって、
 - (a) 該転写調節因子と被験サンプルとを接触させる工程、
 - (b) 該転写調節因子と被験サンプルとの結合活性を検出する工程、
 - (c) 該転写調節因子に結合する活性を有する化合物を選択する工程、を含む方法、
13. (1)から(5)のいずれかに記載の転写調節因子と、hSNF2H、hSNF2L、NCoA-62/Skip およびそれらの他種相同体からなる群より選択されるタンパク質との結合を阻害または促進する化合物をスクリーニングする方法であって、
 - (a) 被検サンプルの存在下で、該転写調節因子と hSNF2H、hSNF2L、NCoA-62/Skip またはそれらの他種相同体とを接触させる工程、
 - (b) 該転写調節因子と hSNF2H、hSNF2L、NCoA-62/Skip またはそれらの他種相

同体との結合活性を検出する工程、

(c) 被検サンプル非存在下で検出した場合(対照)と比較して、該結合活性を低下または増加させる化合物を選択する工程、を含む方法、

14. (13)に記載の方法により単離しうる、(1)から(5)のいずれかに記載の転写調節因子と、hSNF2H、hSNF2L、NCoA-62/Skip およびそれらの他種相同体からなる群より選択されるタンパク質との結合を阻害する化合物、

15. 配列番号:2または9に記載の塩基配列からなるDNAと特異的にハイブリダイズし、少なくとも15ヌクレオチドの鎖長を有するDNA、に関する。

なお、本発明において「転写調節因子」とは、遺伝子の発現を調節しているタンパク質を指す。また、「プロモドメイン」とは、シグナル依存的な転写に関連している転写調節因子中で保存されているタンパク質-タンパク質相互作用に関与するアミノ酸のモチーフを指す。

本発明は、プロモドメインを有する新規な転写調節因子に関する。本発明の転写調節因子に含まれる、「TCoA1」と命名されたタンパク質のアミノ酸配列およびそのバリエーションのアミノ酸配列をそれぞれ配列番号:1および10に、そのcDNAの塩基配列をそれぞれ配列番号:2および9に示す(以下、特に断らない限り、これらを併せて「TCoA1」と称する)。「TCoA1」は1つのコスミドF26H11のゲノム配列から同定された機能が未知の遺伝子である線虫(*C. Elegans*)染色体III遺伝子F26H11.2, F26H11.3a および F26H11.3b (Wilson et al., (1994) *Nature*, 368, 32-38)の推定タンパク質に最も深い関連がある。これら2つの推定される線虫タンパク質と「TCoA1」タンパク質のアミノ酸配列を比較すると、それらはドメインの配置は異なるものの非常に類似したタンパク質である。

多くのプロモドメインタンパク質と同様に、「TCoA1」は一個所のプロモドメインを保持する。TIFファミリー、GCN5とP/CAFの構造と類似して、このドメインはカルボキシ末端に近接して存在する(Jeanmougin et al. (1997). *Trends Biochem. Sci.* 22, 151-153)。「TCoA1」には、他のプロモドメインタンパク質

と同様に、C4HC3 ジンクフィンガーが存在する。プロモドメインと C4HC3 ジンクフィンガーの組み合わせは、これまでいくつかの白血病の遺伝子切断点でしばしば発見されたものである(Tkachuk et al.(1992). Cell, 71, 691-700、Gu et al.(1992) Cell, 71, 701-708、Miki et al.(1991) Proc. Nat. Acad. Sci., 88,5167-5171、Le Douarin et al.(1995) EMBO J. 14, 2020-2033、Borrow et al.(1996). Nature Genet., 14, 33-41)。従って、「TCoA1」は染色体第 17 番 q23 に関与する切断遺伝子の候補である。

「TCoA1」には、核移行シグナルのモチーフが多数存在する。このことは「TCoA1」タンパク質が核内に局在することを示している。「TCoA1」には、他のプロモドメインタンパク質と同様に、核受容体と相互作用する部位を規定すると予測される LXXLL モチーフのシリーズが存在する(Heery et al.(1997). Trends Biochem. Sci. 22, 151-153、Torchia et al.(1997). Nature 387, 677-684)。LXXLL ドメインを介した、リガンドが結合した受容体との相互作用は、転写コアクチベクターとして「TCoA1」が機能することを示唆する。また、

「TCoA1」のカルボキシ末端には非常に広範なグルタミンリッチなドメインが存在する。グルタミンリッチなドメインは、p300/CBP(Shikama et al.(1997) Trends in Cell Biol. 2, 230-236)やショウジョウバエ (drosophila) の fsh (Haynes et al.(1989). Dev. Biol., 134, 246-257)のようなプロモドメインを持つタンパク質を含め、多くの転写調節因子で同定されている。これらの酸性領域は、活性物質としての機能を規定するタンパク-タンパク相互作用に関与することが予測される (Courey et al.(1989) Cell 59, 827-836)。

「TCoA1」タンパク質には、このように癌などの増殖性疾患に関連することが知られている他のプロモドメインタンパク質と共通した多くの特徴を有する。従って、「TCoA1」タンパク質もまた癌に関連していることが考えられ、「TCoA1」タンパク質やその遺伝子、「TCoA1」タンパク質の機能を調節する化合物などは、癌や他の増殖性疾患の治療や予防に応用しうる。

また、「TCoA1」と相互作用する hSNF2H および hSNF2L がクロマチン介在転写制御機構の一連の過程に関与していることから、該タンパク質がクロマチンが介在する転写制御においてなんらかの役割を果たすことを強く示唆するものである。従って、TCoA1 はクロマチン再構成機構に関連することにより、核受容体への転写応答を統合するタンパク質として大きな役割を果たしていることが考えられる。

本発明の転写調節因子は、当業者に公知の方法により、遺伝子組み換え技術を用いて調製される組み換えタンパク質として、また天然のタンパク質として調製することが可能である。組み換えタンパク質であれば、例えば、本発明の転写調節因子をコードする DNA（例えば、配列番号：2 または 9 に記載の塩基配列を有する DNA）を適当な発現ベクターに組み込み、これを宿主細胞に導入して得た形質転換体から精製するなどの方法により調製することが可能である。また、天然のタンパク質であれば、例えば、調製した組み換えタンパク質を小動物に免疫することにより得た抗体を用いたカラムを調製し、本発明の転写調節因子の発現の高い組織や細胞（例えば、精巣や癌細胞など）の抽出物に対し該カラムを用いたアフィニティークロマトグラフィーを行うなどの方法により調製することが可能である。

本発明は、また、「TCoA1」タンパク質（配列番号：1 または 10）と機能的に同等な転写調節因子を包含する。このような転写調節因子には、「TCoA1」タンパク質（配列番号：1 または 10）の変異体や種々の生物由来の「TCoA1」タンパク質が含まれる。

あるタンパク質と機能的に同等なタンパク質を単離する方法としては、タンパク質中のアミノ酸に変異を導入する方法が当業者によく知られている。即ち、当業者であれば、例えば、PCR による部位特異的変異誘発システム（GIBCO-BRL 社, Gaithersburg, Maryland）、オリゴヌクレオチドによる部位特異的変異誘発法（Kramer, W. and Fritz, HJ (1987) *Methods in Enzymol.*, 154:350-367）などの

方法を利用して、配列番号：1 または 10 に記載の「TCoA1」タンパク質の機能に影響を与えないアミノ酸を適宜置換などして、「TCoA1」タンパク質と同等の機能を有する転写調節因子を単離することは通常行いうることである。また、タンパク質中のアミノ酸の変異は自然界においても生じることもある。本発明の転写調節因子には、配列番号：1 または 10 に記載の「TCoA1」タンパク質のアミノ酸配列において 1 もしくは複数のアミノ酸が置換、欠失、付加、および／または挿入されたアミノ酸配列からなり、hSNF2H、hSNF2L や NCoA-62/Skip との結合活性を有するタンパク質、および配列番号：1 または 10 に記載の「TCoA1」タンパク質のアミノ酸配列において 1 もしくは複数のアミノ酸が置換、欠失、付加、および／または挿入されたアミノ酸配列からなり、プロモドメインを保持している転写調節因子が含まれる。

変異するアミノ酸の個数は、「TCoA1」タンパク質の機能が保持される限り特に制限はないが、通常、50 アミノ酸以内であり、好ましくは 30 アミノ酸以内であり、さらに好ましくは、10 アミノ酸以内であり、さらに好ましくは 3 アミノ酸以内である。また、アミノ酸の変異部位も「TCoA1」タンパク質の機能が保持される限り特に制限はない。

あるアミノ酸配列に対する 1 又は複数個のアミノ酸残基の欠失、付加及び／又は他のアミノ酸による置換により修飾されたアミノ酸配列を有するタンパク質がその生物学的活性を維持することはすでに知られている (Mark, D. F. et al., Proc. Natl. Acad. Sci. USA (1984) 81, 5662-5666、Zoller, M. J. & Smith, M. Nucleic Acids Research (1982) 10, 6487-6500、Wang, A. et al., Science 224, 1431-1433、Dalbadie-McFarland, G. et al., Proc. Natl. Acad. Sci. USA (1982) 79, 6409-6413)。

また、変異するアミノ酸残基においては、アミノ酸側鎖の性質が保存されている別のアミノ酸に変異されることが望ましい。例えばアミノ酸側鎖の性質としては、疎水性アミノ酸 (A、I、L、M、F、P、W、Y、V)、親水性アミノ酸 (R、

D、N、C、E、Q、G、H、K、S、T)、脂肪族側鎖を有するアミノ酸 (G、A、V、L、I、P)、水酸基含有側鎖を有するアミノ酸 (S、T、Y)、硫黄原子含有側鎖を有するアミノ酸 (C、M)、カルボン酸及びアミド含有側鎖を有するアミノ酸 (D、N、E、Q)、塩基含有側鎖を有するアミノ酸 (R、K、H)、芳香族含有側鎖を有するアミノ酸 (H、F、Y、W) を挙げることができる (括弧内はいずれもアミノ酸の一字標記を表す)。

「TCoA1」タンパク質のアミノ酸配列 (配列番号: 1 または 10) に複数個のアミノ酸残基が欠失したタンパク質としては、例えば、hSNF2H、hSNF2L、NCoA-62/Skip またはそれらの他種相同体との結合活性を有する「TCoA1」タンパク質の部分ペプチドが含まれる。実施例 6 (図 5) に記載のように「TCoA1」タンパク質の N 末端側アミノ酸残基には、hSNF2H、hSNF2L、NCoA-62/Skip またはそれらの他種相同体との結合活性が存在する。このようなペプチドは、生体内において「TCoA1」タンパク質と上記したその結合タンパク質との結合を阻害し、「TCoA1」タンパク質の生体内での機能を阻害するために利用することが可能である。

「TCoA1」タンパク質のアミノ酸配列 (配列番号: 1 または 10) に複数個のアミノ酸残基が付加されたタンパク質としては、例えば、「TCoA1」タンパク質を含む融合タンパク質が挙げられる。融合タンパク質は、「TCoA1」タンパク質と他のペプチド又はタンパク質とが融合したものであり、本発明に含まれる。融合タンパク質を作製する方法は、本発明の「TCoA1」タンパク質をコードする DNA と他のペプチド又はタンパク質をコードする DNA をフレームが一致するように連結してこれを発現ベクターに導入し、宿主で発現させればよく、当業者に公知の手法を用いることができる。本発明のタンパク質との融合に付される他のペプチド又はタンパク質としては、特に限定されない。

本発明のタンパク質との融合に付される他のペプチドとしては、例えば、FLAG (Hopp, T. P. et al., BioTechnology (1988) 6, 1204-1210)、6 個の His (ヒ

スチジン) 残基からなる 6×His、10×His、インフルエンザ凝集素 (HA)、ヒト c-myc の断片、VSV-GP の断片、p18HIV の断片、T7-tag、HSV-tag、E-tag、SV40T 抗原の断片、lck tag、 α -tubulin の断片、B-tag、Protein C の断片等の公知のペプチドを使用することができる。また、本発明のタンパク質との融合に付される他のタンパク質としては、例えば、GST (グルタチオン-S-トランスフェラーゼ)、HA (インフルエンザ凝集素)、イムノグロブリン定常領域、 β -ガラクトシダーゼ、MBP (マルトース結合タンパク質) 等が挙げられる。

市販されているこれらペプチドまたはタンパク質をコードする DNA を本発明のタンパク質をコードする DNA と融合させ、これにより調製された融合 DNA を発現させることにより、融合タンパク質を調製することができる。

また、あるタンパク質と機能的に同等なタンパク質を単離するための他の方法としては、ハイブリダイゼーション技術 (Sambrook, J et al., Molecular Cloning 2nd ed. 9.47-9.58, Cold Spring Harbor Lab. press, 1989) を利用する方法が当業者によく知られている。即ち、当業者にとっては、ハイブリダイゼーション技術 (Sambrook, J et al., Molecular Cloning 2nd ed. 9.47-9.58, Cold Spring Harbor Lab. press, 1989) を用いて、配列番号: 2 に記載の「TCoA1」タンパク質をコードする DNA 配列もしくはその一部を基に、これと相同性の高い DNA を単離して、該 DNA から「TCoA1」タンパク質と機能的に同等な転写調節因子を得ることも通常行いうることである。本発明の転写調節因子には、配列番号: 2 に記載の「TCoA1」タンパク質をコードする DNA とハイブリダイズする DNA がコードし、プロモドメインを有する転写調節因子も含まれる。機能的に同等な転写調節因子を単離するための生物としては、ヒト以外に、例えば、マウス、ラット、ウシ、サル、ブタなどが挙げられるが、これらに制限されない。機能的に同等な転写調節因子をコードする DNA を単離するためのハイブリダイゼーションのストリンジエンシーは、通常、「42°C、2xSSC、0.1% SDS」程度であり、好ましくは「50°C、2xSSC、0.1% SDS」程度、さらに好ましくは「65°C、2xSSC、

0.1% SDS」程度と、温度を上げる程に高い相同性を有する DNA を得ることができる。但し、ハイブリダイゼーションのストリンジェンシーに影響する要素としては温度以外にも複数の要素が考えられ、当業者であればこれら要素を適宜選択することで同様のストリンジェンシーを実現することが可能である。

また、ハイブリダイゼーションにかえて、「TCoA1」タンパク質をコードする配列番号：9 に記載の DNA 配列の配列情報を基に合成したプライマーを用いる遺伝子増幅法、例えば、ポリメラーゼ連鎖反応（PCR）法を利用して、「TCoA1」タンパク質と機能的に同等な転写調節因子をコードする DNA を単離することも可能である。

これらハイブリダイゼーション技術または遺伝子増幅技術により単離される DNA がコードするタンパク質は、通常、「TCoA1」タンパク質とアミノ酸配列において高い相同性を有する。高い相同性とは、「TCoA1」タンパク質とアミノ酸配列において、通常、40%以上の相同性、好ましくは 60%以上の相同性、さらに好ましくは 80%以上の相同性、さらに好ましくは 95%以上の相同性を指す。タンパク質の相同性を決定するには、文献 (Wilbur, W. J. and Lipman, D. J. Proc. Natl. Acad. Sci. USA (1983) 80, 726-730) に記載のアルゴリズムにしたがえばよい。

このようなハイブリダイゼーション技術や遺伝子増幅技術を利用して単離される TcoA1 タンパク質（配列番号：1 または 10）と機能的に同等な転写調節因子としては、TcoA1 タンパク質（配列番号：1 または 10）と一次構造において高い相同性を有し、hSNF2H、hSNF2L や NCoA-62/Skip との結合活性を有する転写調節因子、および TcoA1 タンパク質（配列番号：10）と一次構造において高い相同性を有し、癌に関連した機能の上で重要であると考えられるモチーフであるプロモドメインを保持している転写調節因子が含まれる。

これら転写調節因子は、プロモドメイン以外にも、他のタンパク質との相互作用に関与する配列（例えば、ロイシンジッパー、LXXLL モチーフ）、DNA との

結合などに関与する配列（例えば、ジンクフィンガー）、核移行シグナルなどを有しうる。

タンパク質中のプロモドメインの存在は、DNASIS（日立ソフトウェアエンジニアリング社）上のプロモドメインモチーフ PROSITE データベースを検索することにより決定することが可能である。

また、本発明は、本発明の転写調節因子をコードする DNA に関する。本発明の DNA としては、本発明の転写調節因子をコードしうるものであれば特に制限はなく、cDNA、ゲノム DNA、化学合成 DNA などが含まれる。また、本発明のタンパク質をコードしうる限り、遺伝暗号の縮重に基づく任意の塩基配列を有する DNA が含まれる。本発明のタンパク質をコードする cDNA は、例えば、本発明の転写調節因子をコードする DNA の配列情報（例えば、配列番号：9）を基にプライマーを調製し、ブランク PCR を行うことにより調製することが可能である（例えば、文献「Affara NA et al (1994) Genomics, 22, 205-210」参照）。また、ゲノム DNA であれば、例えば、市販の「Qiagen genomic DNA kits」（Qiagen 社, Hilden, Germany）を用いた方法により調製することが可能である。得られた DNA の塩基配列は、市販の「dye terminator sequencing kit」（Applied Biosystems 社）などを用いて常法により決定することが可能である。本発明の DNA は、後述するように組み換えタンパク質の製造に用いられる他、遺伝子治療などへの応用が考えられる。

また、本発明は、本発明の DNA が挿入されたベクターに関する。本発明の DNA が挿入されるベクターとしては特に制限はなく、本発明の転写調節因子を生体内で発現させるためのベクター、組み換えタンパク質を調製するためのベクターなど目的に応じて種々のベクターが用いられる。本発明の転写調節因子を生体内で発現させるため（特に遺伝子治療のため）に用いられるベクターとしては、各種ウイルスベクターや非ウイルスベクターが挙げられる。ウイルスベクターとしては、例えば、アデノウイルスベクター「pAdexLcw」やレトロウイル

スベクター「pZIPneo」などが、非ウイルスベクターとしては、例えば、カチオニックリポソームなどが挙げられる。本発明の転写調節因子を生産する目的においてベクターを使用する場合には、特に、発現ベクターが有用である。発現ベクターとしては、例えば、大腸菌 (*E. coli*) を用いる場合には「pREP4」(Qiagen 社, Hilden, Germany) などが、酵母を用いる場合には「SP-Q01」(Stratagene 社, La Jolla, California) などが、昆虫細胞を用いる場合には「BAC-to-BAC baculovirus expression system」(GIBCO-BRL 社, Gaithersburg, Maryland) などが好適であるが、これらに制限されない。また、哺乳動物細胞、例えば、CHO 細胞、COS 細胞、NIH3T3 細胞などを用いる場合には、例えば、「LacSwitch II expression system」(Stratagene 社, La Jolla, California) などが好適であるが、これらに制限されない。ベクターへの本発明の DNA の挿入は、常法により行うことができる。

また、本発明は、本発明の DNA を発現可能に保持する形質転換体に関する。本発明の形質転換体には、本発明の DNA が挿入された上記ベクターを保持するもの、本発明の DNA が宿主ゲノム内に組み込まれているものなどが含まれるが、本発明の DNA を発現可能に保持している限り、その存在形態は問わない。本発明のベクターが導入される細胞としては特に制限はない。例えば、遺伝子治療目的の場合には、疾患に応じて、種々の細胞を標的細胞として用いることが可能である。また、本発明の転写調節因子を製造する目的の場合には、例えば、大腸菌、酵母、動物細胞、昆虫細胞などを宿主として用いることが可能である。細胞へのベクターの導入は、例えば、電氣的穿孔法、熱ショックなどの公知の方法で行うことが可能である。

組み換えタンパク質を製造するために作製した形質転換体からの該組み換えタンパク質の分離、精製は、常法により行うことが可能である。例えば、形質転換体を回収し、抽出物を得た後、イオン交換、逆相、ゲル濾過などのクロマトグラフィー、あるいは本発明のタンパク質に対する抗体をカラムに固定した

アフィニティークロマトグラフィーにかけることにより、または、さらにこれらのカラムを複数組み合わせることにより目的のタンパク質を精製し、調製することが可能である。

また、本発明のタンパク質をグルタチオン S トランスフェラーゼタンパク質との融合タンパク質として、あるいはヒスチジンを複数付加させた組み換えタンパク質として宿主細胞（例えば、動物細胞や大腸菌など）内で発現させた場合には、発現させた組み換えタンパク質はグルタチオンカラムあるいはニッケルカラムを用いて精製することができる。融合タンパク質の精製後、必要に応じて融合タンパク質のうち目的のタンパク質以外の領域を、トロンビンまたはファクターXa などにより切断し、除去することも可能である。

また、本発明は、本発明の転写調節因子に結合する抗体に関する。本発明の抗体の形態には、特に制限はなく、ポリクローナル抗体の他、モノクローナル抗体も含まれる。また、ウサギなどに本発明の転写調節因子を免疫して得た抗血清、すべてのクラスのポリクローナル抗体およびモノクローナル抗体、さらに遺伝子組み換えによるヒト型化抗体、ヒト抗体も含まれる。本発明の抗体は、以下の方法により調製することが可能である。ポリクローナル抗体であれば、例えば、本発明の転写調節因子をウサギなどの小動物に免疫し血清を得て、これを本発明の転写調節因子をカップリングさせたアフィニティークラムにより、本発明の転写調節因子のみを認識する画分を得て、さらにこの画分から免疫グロブリン G あるいは M を、プロテイン A、あるいはプロテイン G カラムにより精製することにより調製することができる。また、モノクローナル抗体であれば、本発明の転写調節因子をマウスなどの小動物に免疫を行い、同マウスより脾臓を摘出し、これをすりつぶして細胞にし、マウスミエローマ細胞とポリエチレングリコールなどの試薬により融合させ、これによりできた融合細胞（ハイブリドーマ）の中から、本発明の転写調節因子に対する抗体を産生するクローンを選択する。次いで、得られたハイブリドーマをマウス腹腔内に移植し、

同マウスより腹水を回収し、得られたモノクローナル抗体を、例えば、硫酸沈殿、プロテイン A、プロテイン G カラム、DEAE イオン交換クロマトグラフィー、本発明の転写調節因子をカップリングしたアフィニティーカラムなどにより精製することで調製することが可能である。本発明の抗体は、本発明の転写調節因子の精製や検出に用いられる他、本発明の転写調節因子の機能を抑制するための薬剤として用いることも可能である。抗体を薬剤として用いる場合には、免疫原性の点で、ヒト抗体またはヒト化抗体が有効である。ヒト抗体またはヒト化抗体は当業者に公知の方法により調製することができる。例えば、ヒト抗体は、免疫系をヒトと入れ換えたマウスに本発明の転写調節因子を免疫することにより調製することが可能である。また、ヒト化抗体は、モノクローナル抗体産生細胞から抗体遺伝子をクローニングし、その抗原決定部位を既存のヒト抗体に移植する CDR グラフト法により調製することが可能である。

また、本発明は、本発明の転写調節因子に結合する化合物のスクリーニング方法に関する。本発明のスクリーニング法は、本発明の転写調節因子と被験サンプルとを接触させ、被験サンプルと本発明の転写調節因子との結合活性を検出し、本発明の転写調節因子に結合する活性を有する化合物を選択する工程を含む。スクリーニングに用いる被験サンプルとしては特に制限はなく、例えば、細胞抽出液、細胞培養上清、合成低分子化合物のライブラリー、精製タンパク質、遺伝子ライブラリーの発現産物、合成ペプチドのライブラリーなどが挙げられる。

本発明の転写調節因子を用いてこれに結合する化合物を単離する方法としては、例えば、以下の方法が当業者によく知られている。本発明の転写調節因子に結合するタンパク質のスクリーニングは、例えば、本発明の転写調節因子と結合するタンパク質を発現していることが予想される組織若しくは細胞(例えば、精巣組織細胞や癌細胞系など)よりファージベクター (λ gt11, ZAP など) を用いた cDNA ライブラリーを作製し、これを LB-アガロース上で発現させフィル

ターに発現させたタンパク質を固定し、本発明の転写調節因子をビオチンラベル、あるいは GST タンパク質との融合タンパク質として精製し、これを上記フィルターと反応させ、結合するタンパク質を発現しているブランクを、ストレプトアビジン、あるいは抗 GST 抗体により検出する「ウエストウエスタンブロッティング法」(Skolnik EY, Margolis B, Mohammadi M, Lowenstein E, Fischer R, Drepps A, Ullrich A, and Schlessinger J (1991)Cloning of PI3 kinase-associated p85 utilizing a novel method for expression/cloning of target proteins for receptor tyrosine kinases. Cell 65, 83-90) により実施することが可能である。また、本発明の転写調節因子に結合するタンパク質またはその遺伝子のスクリーニングは、「2 ハイブリッドシステム」(「MATCHMARKER Two-Hybrid System」,「Mammalian MATCHMAKER Two-Hybrid Assay Kit」,「MATCHMAKER One-Hybrid System」(いずれも clontech 社)、「HybriZAP Two-Hybrid Vector System」(stratagene 社)、文献「Dalton S, and Treisman R (1992) Characterization of SAP-1, a protein recruited by serum response factor to the c-fos serum response element. Cell 68, 597-612」) に従い実施することも可能である。2 ハイブリッドシステムにおいては、本発明の転写調節因子を SRF 結合領域または GAL4 結合領域と融合させて酵母細胞の中で発現させ、本発明の転写調節因子と結合するタンパク質を発現していることが予想される細胞より、VP16 または GAL4 転写活性化領域と融合する形で発現するような cDNA ライブラリーを作製する。これを上記酵母細胞に導入し、検出された陽性クローンからライブラリー由来 cDNA を単離する(酵母細胞内で本発明の転写調節因子と結合するタンパク質が発現すると、両者の結合によりレポーター遺伝子が活性化され、陽性のクローンが確認できる)。さらに、該 cDNA を大腸菌に導入して、該 cDNA によりコードされるタンパク質を発現させ、該タンパク質を回収することができる。

本発明の転写調節因子に結合するタンパク質のスクリーニングは、さらに、

本発明の転写調節因子を固定したアフィニティーカラムに本発明の転写調節因子と結合するタンパク質を発現していることが予想される細胞の培養上清もしくは細胞抽出物をのせ、カラムに特異的に結合するタンパク質を精製することにより実施することも可能である。

また、固定した本発明の転写調節因子に、合成化合物、または天然物バンク、もしくはランダムファージペプチドディスプレイライブラリーを作用させ、結合する分子をスクリーニングする方法や、コンビナトリアルケミストリー技術によるハイスループットを用いたスクリーニング (Wrighton NC; Farrell FX; Chang R; Kashyap AK; Barbone FP; Mulcahy LS; Johnson DL; Barrett RW; Jolliffe LK; Dower WJ., Small peptides as potent mimetics of the protein hormone erythropoietin, Science (UNITED STATES) Jul 26 1996, 273 p458-64, Verdine GL., The combinatorial chemistry of nature. Nature (ENGLAND) Nov 7 1996, 384 p11-13, Hogan JC Jr., Directed combinatorial chemistry. Nature (ENGLAND) Nov 7 1996, 384 p17-9) により本発明の転写調節因子に結合する、低分子化合物、タンパク質 (またはその遺伝子)、ペプチドなどを単離する方法も当業者に周知の技術である。

本発明において、結合した化合物を検出又は測定する手段として表面プラズモン共鳴現象を利用したバイオセンサーを使用することもできる。表面プラズモン共鳴現象を利用したバイオセンサーは本発明のタンパク質と被検化合物との間の相互作用を微量のタンパク質を用いてかつ標識することなく、表面プラズモン共鳴シグナルとしてリアルタイムに観察することが可能である (例えば BIAcore、Pharmacia 製)。したがって、BIAcore 等のバイオセンサーを用いることにより本発明のタンパク質と被験化合物との結合を評価することが可能である。

また、本発明は、本発明の転写調節因子とこれに相互作用するタンパク質との結合を阻害または促進する化合物のスクリーニング方法に関する。本発明に

において、TcoA1 タンパク質と、hSNF2H、hSNF2L および NCoA-62/Skip との結合が検出されたことに基づき、このようなスクリーニングを行なうことが可能となった。このスクリーニングは、(a) 被検サンプルの存在下で、本発明の転写調節因子と hSNF2H、hSNF2L、NCoA-62/Skip またはそれらの他種相同体とを接触させる工程、(b) 本発明の転写調節因子と hSNF2H、hSNF2L、NCoA-62/Skip またはそれらの他種相同体との結合活性を検出する工程、および (c) 被検サンプル非存在下で検出した場合（対照）と比較して、該結合活性を低下させる化合物を選択する工程、を含む方法により実施することが可能である。

被検サンプルとしては特に制限はなく、例えば、細胞抽出液、細胞培養上清、合成低分子化合物のライブラリー、精製タンパク質、遺伝子ライブラリーの発現産物、合成ペプチドのライブラリーなどが挙げられる。また、上記の本発明のタンパク質に結合する化合物のスクリーニングにより単離された化合物を被検試料として用いることも可能である。

また、スクリーニングに用いる本発明の転写調節因子としては、完全なタンパク質であっても、hSNF2H、hSNF2L、NCoA-62/Skip またはそれらの他種相同体との結合領域を含む部分ペプチドであってもよい。また、スクリーニングに用いる hSNF2H、hSNF2L、NCoA-62/Skip またはそれらの他種相同体としては、完全なタンパク質であってもよく、また、本発明の転写調節因子との結合領域を含む部分ペプチドであってもよい。

本発明の転写調節因子と hSNF2H、hSNF2L、NCoA-62/Skip またはそれらの他種相同体との結合活性の検出は、例えば、以下のように行なうことができる。

マイクロプレートに固定した本発明の転写調節因子に被検サンプルと hSNF2H、hSNF2L、NCoA-62/Skip またはそれらの他種相同体を存在せしめ、hSNF2H、hSNF2L、NCoA-62/Skip またはそれらの他種相同体 に対するマウスやウサギ抗体を反応させ、さらにペルオキシダーゼやアルカリフォスファターゼなどで標識した抗マウスあるいは抗ウサギ抗体を加えて反応させた後、標識した酵素の

基質を添加し、その酵素活性を測定する。この酵素活性が、被検サンプルを添加しなかった場合において検出した酵素活性と比較して、低いまたは高い値を示すような化合物を選択する。これにより、本発明の転写調節因子と hSNF2H、hSNF2L、NCoA-62/Skip またはそれらの他種相同体との間の結合を阻害または促進する活性を有する化合物を得ることができる。

このスクリーニングは、固定するタンパク質を hSNF2H、hSNF2L、NCoA-62/Skip またはそれらの他種相同体として、被験試料とともに加えるタンパク質を本発明の転写調節因子としてもよい。

また、被験サンプルと加える本発明の転写調節因子あるいは hSNF2H、hSNF2L、NCoA-62/Skip またはそれらの他種相同体を直接にペルオキシダーゼやアルカリフォスファターゼで標識したり、これらの酵素との融合タンパク質としたものを使用してもよい。また、これらの酵素に限らず、ルシフェラーゼやベータガラクトシダーゼ、GFP タンパク質、などとの融合タンパク質として発現し、被験サンプルによるそれらの酵素活性の阻害あるいは促進を測定することにより本発明の転写調節因子と hSNF2H、hSNF2L、NCoA-62/Skip またはそれらの他種相同体との間の結合を阻害または活性化する活性を有する化合物を選択することができる。

また、本発明の転写調節因子とこれに相互作用するタンパク質との結合を阻害または促進する化合物のスクリーニングは、哺乳動物の 2 ハイブリッドシステム (Clontech, Palo Alto) を利用して行なうことができる。即ち、2 ハイブリッドシステムを利用して、哺乳動物細胞内で本発明の転写調節因子とこれに相互作用するタンパク質とを発現させ、該哺乳動物細胞に被検サンプルを添加して、その後、レポーター活性を検出する。検出されたレポーター活性が、被検サンプルを添加しなかった場合において検出したレポーター活性と比較して、低いまたは高い値を示すような化合物を選択する。これにより、本発明の転写調節因子と hSNF2H、hSNF2L、NCoA-62/Skip またはそれらの他種相同体 との間

の結合を阻害または促進する活性を有する化合物を得ることができる。

本発明のスクリーニングにより単離される化合物は、癌および他の増殖性疾患の治療や予防への応用が考えられる。単離される化合物をヒトや哺乳動物、例えばマウス、ラット、モルモット、ウサギ、ニワトリ、ネコ、イヌ、ヒツジ、ブタ、ウシ、サル、マントヒヒ、チンパンジーの医薬として使用する場合には、単離された化合物自体を直接患者に投与する以外に、公知の製剤学的方法により製剤化して投与を行うことも可能である。例えば、必要に応じて糖衣を施した錠剤、カプセル剤、エリキシル剤、マイクロカプセル剤として経口的に、あるいは水もしくはそれ以外の薬学的に許容し得る液との無菌性溶液、又は懸濁液剤の注射剤の形で非経口的に使用できる。例えば、薬理学上許容される担体もしくは媒体、具体的には、滅菌水や生理食塩水、植物油、乳化剤、懸濁剤、界面活性剤、安定剤、香味剤、賦形剤、ベヒクル、防腐剤、結合剤などと適宜組み合わせ、一般に認められた製薬実施に要求される単位用量形態で混和することによって製剤化することが考えられる。これら製剤における有効成分量は指示された範囲の適当な容量が得られるようにするものである。

錠剤、カプセル剤に混和することができる添加剤としては、例えばゼラチン、コーンスターチ、トラガントガム、アラビアゴムのような結合剤、結晶性セルロースのような賦形剤、コーンスターチ、ゼラチン、アルギン酸のような膨化剤、ステアリン酸マグネシウムのような潤滑剤、シヨ糖、乳糖又はサッカリンのような甘味剤、ペパーミント、アカモノ油又はチェリーのような香味剤が用いられる。調剤単位形態がカプセルである場合には、上記の材料にさらに油脂のような液状担体を含有することができる。注射のための無菌組成物は注射用蒸留水のようなベヒクルを用いて通常の製剤実施に従って処方することができる。

注射用の水溶液としては、例えば生理食塩水、ブドウ糖やその他の補助薬を含む等張液、例えば D-ソルビトール、D-マンノース、D-マンニトール、塩化ナ

トリウムが挙げられ、適当な溶解補助剤、例えばアルコール、具体的にはエタノール、ポリアルコール、例えばプロピレングリコール、ポリエチレングリコール、非イオン性界面活性剤、例えばポリソルベート 80 (TM)、HCO-50 と併用してもよい。

油性液としてはゴマ油、大豆油があげられ、溶解補助剤として安息香酸ベンジル、ベンジルアルコールと併用してもよい。また、緩衝剤、例えばリン酸塩緩衝液、酢酸ナトリウム緩衝液、無痛化剤、例えば、塩酸プロカイン、安定剤、例えばベンジルアルコール、フェノール、酸化防止剤と配合してもよい。調製された注射液は通常、適当なアンプルに充填させる。

患者への投与は、例えば、動脈内注射、静脈内注射、皮下注射などのほか、鼻腔内的、経気管支的、筋内的、または経口的に当業者に公知の方法により行いうる。投与量は、患者の体重や年齢、投与方法などにより変動するが、当業者であれば適当な投与量を適宜選択することが可能である。また、該化合物が DNA によりコードされうるものであれば、該 DNA を遺伝子治療用ベクターに組み込み、遺伝子治療を行うことも考えられる。投与量、投与方法は、患者の体重や年齢、症状などにより変動するが、当業者であれば適宜選択することが可能である。

例えば、本発明のタンパク質と結合する化合物や本発明のタンパク質の活性を調節する化合物の投与量は、症状により差異はあるが、経口投与の場合、一般的に成人（体重 60kg として）においては、1 日あたり約 0.1 から 100mg、好ましくは約 1.0 から 50mg、より好ましくは約 1.0 から 20mg である。

非経口的に投与する場合は、その 1 回投与量は投与対象、対象臓器、症状、投与方法によっても異なるが、例えば注射剤の形では通常成人（体重 60kg として）においては、1 日あたり約 0.01 から 30mg、好ましくは約 0.1 から 20mg、より好ましくは約 0.1 から 10mg 程度を静脈注射により投与するのが好都合である。他の動物の場合も、体重 60kg 当たりに換算した量を投与することができる。

また、本発明は、「TCoA1」タンパク質をコードする DNA と特異的にハイブリダイズし、少なくとも 15 ヌクレオチドの鎖長を有する DNA に関する。ここで「特異的にハイブリダイズする」とは、通常のハイブリダイゼーション条件下、好ましくはストリンジェントなハイブリダイゼーション条件下で、他のタンパク質をコードする DNA とクロスハイブリダイゼーションが有意に生じないことを指す。このような DNA には、本発明のタンパク質をコードする DNA 又は該 DNA と相補的な DNA と特異的にハイブリダイズし得るプローブやプライマー、ヌクレオチド又はヌクレオチド誘導体（例えばアンチセンスオリゴヌクレオチドやリボザイム等）が含まれる。

本発明は、例えば、配列番号：2 または 9 に示される塩基配列中のいずれかの箇所にハイブリダイズするアンチセンスオリゴヌクレオチドを含む。このアンチセンスオリゴヌクレオチドは、好ましくは配列番号：2 または 9 に示される塩基配列中の連続する少なくとも 15 個以上のヌクレオチドに対するアンチセンスオリゴヌクレオチドである。さらに好ましくは、前記連続する少なくとも 15 個以上のヌクレオチドが翻訳開始コドンを含む、前記のアンチセンスオリゴヌクレオチドである。

アンチセンスオリゴヌクレオチドとしては、それらの誘導体や修飾体を使用することができる。このような修飾体として、例えばメチルホスホネート型又はエチルホスホネート型のような低級アルキルホスホネート修飾体、ホスホロチオエート修飾体又はホスホロアミデート修飾体等が挙げられる。

ここでいう「アンチセンスオリゴヌクレオチド」とは、DNA 又は mRNA の所定の領域を構成するヌクレオチドに対応するヌクレオチドが全て相補的であるもののみならず、DNA または mRNA とオリゴヌクレオチドとが配列番号：9 に示される塩基配列に特異的にハイブリダイズできる限り、1 又は複数個のヌクレオチドのミスマッチが存在していてもよい。

このような DNA としては、少なくとも 15 個の連続したヌクレオチド配列領域

で、少なくとも 70%、好ましくは少なくとも 80%、より好ましくは 90%、さらに好ましくは 95% 以上の塩基配列上の相同性を有するものを示す。なお、相同性を決定するためのアルゴリズムは本明細書に記載したものを使用すればよい。このような DNA は、後述の実施例に記載するように本発明のタンパク質をコードする DNA を検出若しくは単離するためのプローブとして、又は増幅するためのプライマーとして有用である。

本発明のアンチセンスオリゴヌクレオチド誘導体は、本発明のタンパク質の産生細胞に作用して、該タンパク質をコードする DNA 又は mRNA に結合することにより、その転写又は翻訳を阻害したり、mRNA の分解を促進したりして、本発明のタンパク質の発現を抑制することにより、結果的に本発明のタンパク質の作用を抑制する効果を有する。

本発明のアンチセンスオリゴヌクレオチド誘導体は、それらに対して不活性な適当な基剤と混和して塗布剤、パップ剤等の外用剤とすることができる。

また、必要に応じて、賦形剤、等張化剤、溶解補助剤、安定化剤、防腐剤、無痛化剤等を加えて錠剤、散財、顆粒剤、カプセル剤、リボソームカプセル剤、注射剤、液剤、点鼻剤など、さらに凍結乾燥剤とすることができる。これらは常法にしたがって調製することができる。

本発明のアンチセンスオリゴヌクレオチド誘導体は患者の患部に直接適用するか、又は血管内に投与するなどして結果的に患部に到達し得るように患者に適用する。さらには、持続性、膜透過性を高めるアンチセンス封入素材を用いることもできる。例えば、リボソーム、ポリ-L- リジン、リビッド、コレステロール、リボフェクチン又はこれらの誘導体が挙げられる。

本発明のアンチセンスオリゴヌクレオチド誘導体の投与量は、患者の状態に応じて適宜調整し、好ましい量を用いることができる。例えば、0.1 ~100mg/kg、好ましくは 0.1 ~50mg/kg の範囲で投与することができる。

本発明のアンチセンスオリゴヌクレオチドは本発明のタンパク質の発現を阻

害し、従って本発明のタンパク質の生物学的活性を抑制することにおいて有用である。また、本発明のアンチセンスオリゴヌクレオチドを含有する発現阻害剤は、本発明のタンパク質の生物学的活性を抑制することが可能である点で有用である。

図面の簡単な説明

図1は、「TCoA1」に同定されたドメインの配置を示す。図中の記号は、下記を示す。

「CH4C3」；CH4C3 zinc フィンガー、「bHLH」；塩基性ヘリックスループヘリックス、「Q-rich」；グルタミンリッチ、「C2HC4」；C2HC4 zinc フィンガー、「BDM」；プロモドメイン、「↑」；LXXLL モチーフ。

図2は、Aは、染色体第17番に対するモノクロモソームハイブリットセルパネルをプライマーnb15Gとプライマーnb15Hを用いて解析した結果を示す。各々のハイブリッドを含むヒト染色体を示す。133bpの産物はヒト染色体第17番のモノクロモソームであるGM10498細胞系で特異的に増幅された。

Bは、染色体第17番上にある「TCoA1」の位置をGenebridge4放射ハイブリッドパネル解析によって決定した結果を示す。

図3は、「TCoA1」の正常組織における発現のノーザンプロット解析を示す電気泳動写真である。Aは「TCoA1」、Bはアクチンをプローブにフィルターをハイブリダイズした。また図の右側はサイズマーカーを示している。

図4．哺乳動物の2ハイブリッド解析系を用いて、TCoA1とhSNF2H、hSNF2L、およびNCoA-62/Skipとの相互作用を検出した結果を示す。

図5は、TCoA1のC末端領域とNCoA62/SKIP、hSNF2H、またはhSNF2Lとの相互作用の地図を示す。

図6は、TCoA1とNCoA62/SKIPとの相互作用地図を示す。相互作用最小領域(224-317位)を図下に示した。

図7は、TCOA1とhSNF2Hとの相互作用地図を示す。hSNF2Hの相互作用最小領域(921-1017位)を図下に示した。

図8は、TcoA1との相互作用に関連するタンパク質を示す。確認されていない相互作用を波線で示した。

発明を実施するための最良の形態

以下に本発明を実施例によりさらに詳細に説明するが本発明はこれら実施例に制限されるものではない。

[実施例1] 「TCOA1」遺伝子の単離

(1) 新規プロモドメインを有する遺伝子の同定

既知のプロモドメインモチーフをコードする種々の塩基配列を用い、ESTデータベースに対してBLAST検索を行った。テトラヒメナ サーモフィラ HAT A1 遺伝子(Brownwell et al,(1996) Cell 84, 843-851)の塩基配列を用いた検索結果から、プロモドメイン遺伝子をコードする可能性のある幾つかのESTを同定した。その内の一つ、胎児肺 cDNA ライブラリー由来のEST(W17142)は、新規遺伝子を提供することが判明した。

(2) 全長塩基配列の単離

ESTW17142に対する全長cDNAのクローニングを以下のように行った。まずPCRプライマーnb15U (GGATTATGAGGGGTTGAAGAGGG/配列番号:3)およびプライマーnb15L (AAGGCAACAGAGTCTGTAGCCCAA/配列番号:4)をデザインし、精巢のcDNAを鋳型にポリメラーゼ連鎖反応を行い、119bpの産物を得た。PCR産物は直接QIAquick(Qiagen社)の精製カラムを用いて精製した。このPCR産物をプローブとして、精巢cDNAライブラリー(Clontech社 HL3024a)のスクリーニングを行い、さらに上記ESTの配列をもつcDNAクローンを用いて、ライブラリーのスクリーニングを再度行った。なお、プローブはランダムプライミングにより [α - 32 P] dCTPで標識し、クロマスピン10カラム(Clontech社)で精製した。ライ

ブラリーフィルターのハイブリダイゼーションは、ExpressHyb ハイブリダイゼーション溶液(Clontech 社)を用い、65°Cで1時間かけて行った。フィルターを65°Cで0.5x SSC、0.1% SDS の最終的ストリンジェンシーになるまで洗浄した。続いてそのフィルターをシグナル強度増大のために-70°Cで1-3日間オートラジオグラフィーを行った。同様な操作を、得られたクローンを連結させ、本遺伝子の全コード領域をカバーする塩基配列が得られるまで繰り返し行った。全ての塩基配列は、ABI 色素ターミネーター化学を用いて ABI377 自動配列決定機で決定した。5' 端のクローンは非常に GC 含有率が高かったため、配列決定の前にプラスミドにサブクローンした。

ライブラリーのスクリーニングの結果、全体で 9865bp の塩基配列が得られた。この全塩基配列中には、8979 番目の塩基で終止する 2993 アミノ酸をコードするオープンリーディングフレーム(ORF)が存在し、それに続いてポリ A テールまで 877 bp の 3' UTR が存在した(図 1)。この配列は、配列の長さがノーザン分析で示された 10.5kb に匹敵すること、および 5' 端の配列は GC リッチであり、多くの遺伝子の開始点に見られる CpG アイランドが存在することと一致すること (Cross and Bird,(1995) Curr. Opin. Genet. Dev .5,309-314)から、ほぼ完全長であると考えられる。単離された cDNA の塩基配列を配列番号:2 に、該 cDNA がコードするタンパク質のアミノ酸配列を配列番号:1 に示す。

(3) 相同性および転写因子のモチーフ特性の同定

モチーフは PROSITE により検索した。タンパク質の比較は GCG 中の Bestfit を利用して行った。核移行シグナルの同定は PSORT で行った。モチーフ検索の結果、いくつかの保存された領域やドメインが、予測されたタンパク質のアミノ酸配列中に存在していることが判明した(図 1)。これらの保存された部分には、C4HC3 ジンクフィンガー (Aasland et al.(1995). Trends Biochem. Sci., 20, 56-59、Koken et al.(1995). CR Acad. Sci. III, 318, 733-739)、塩基性のヘリックスループヘリックスドメイン (Murre et al. (1989). Cell, 58,

537-544)、広範な疎水性グルタミンリッチなドメイン、CH₂CH₃ ジンクフィンガーとプロモドメインが存在した。さらに、核受容体との相互作用を与えている可能性があるとして予測できる 3 つの LXXLL モチーフ (Torchia et al.(1997). Nature 387, 677-684、Heery et al.(1997) Nature 387, 733-736) が存在した。これら全てのモチーフは、転写調節因子としての機能を示す特性を持つ。PSPORT プログラムの結果、この機能と密接に関連した核位置のコンセンサス配列 (Robbins et al. (1991) Cell, 64, 615-23) が全体で 8箇所存在することが判明した。遺伝子の機能を表すようにこの遺伝子を「TCoA1」(transcriptional co-activator) と命名した。

「TCoA1」の塩基配列を non-redundant DNA データベース上で解析すると、「TCoA1」は 810 残基のタンパク質をコードすると予測される FAC1 遺伝子 (Zhu and Bowser,(1996) Biochemica et Biophysica Acta 1309, 5-8) と 2,183bp において 100%の相同性があることが判明した。FAC1 は元来モノクローナル抗体 Alz50 (Bowser et al. (1995) Dev. Neuroscience 17, 20-37)を用いた発現ライブラリーの免疫ースクリーニングによって単離されている。FAC1 は「TCoA1」とは塩基配列において広範な領域が一致することに加え、転写サイズ (Bowser et al. (1995) Dev. Neuroscience 17, 20-37)や局在 (Zhu and Bowser.(1996) Biochemica et Biophysica Acta 1309, 5-8)においても、FAC1 と重なる領域の外側の塩基配列を用いて得られた「TCoA1」の結果と一致していた。即ち、FAC1 の 2673bp 塩基配列は「TCoA1」の 5'末端 248 から 2631 塩基に相当する部分配列であると考えられる。そこで「TCoA1」の塩基配列と FAC1 塩基配列を比較した結果、一塩基欠失のエラー配列 (2400 の地点で A) が FAC1 に存在することが判明し、これにより ORF の読み枠のシフトと共に翻訳が早期に終止することが想定された。同様に FAC1 の 5'末端配列のエラーによって、メチオニン残基の開始点の誤認識を引き起こしていた。

「TCoA1」の予測されるアミノ酸配列には、線虫 (*C. elegans*) で推定される

タンパク質 F26H11.2, F26H11.3a and F26H11.3b (Wilson et al. (1994). *Nature*, 368, 32-38)に相同ないくつかの広範な領域があった。「Gene Finder」ソフトウェアを用いた解析の結果、F26H11 コスミドに含まれるゲノム配列を検索することにより、これらのタンパク質をコードする遺伝子の予測が可能であった。

「TCoA1」の N 末端は F26H11.g に、また C 末端は F26H11.i に塩基配列の一致する部分が存在した。この結果から「TCoA1」と FAC1 により予測されるこれら 2 つのタンパク質がいずれも線虫の単一のタンパク質に相当することが示唆され、「TCoA1」はこのヒトのホモログであると考えられる。

[実施例 2] 「TCoA1」の染色体地図作製

「TCoA1」の染色体上の位置を決定するために、PCR プライマー nb15G (CCTCAGCTGCAACAAGTCC / 配列番号 : 5) と プライマー nb15H (GCACTGCTTTGCTGAATTTGGA / 配列番号 : 6) を用いて、Coriell Cell Repositories, New Jersey から入手した 24 monochromosomal human/rodent somatic cell lines (Dubois and Naylor. (1993) *Genomics*, 16, 315-319)のそれぞれから得られる DNA を増幅した。その結果、予測通りの 133bp の PCR 産物が細胞系の GM 10567 から増幅され、本遺伝子はヒト第 17 番染色体に存在する可能性を示唆するものであった(図 2A)。

また、「TCoA1」の領域座は GeneBridge4 放射パネル 91 ハイブリッド(Walter et al. (1994). *Nature Genetics*, 7, 22-28) を用いて決定した。再度プライマー G とプライマー H を使い、そのハイブリッドパネルで PCR を行うことによってスクリーニングした。増幅に対して正あるいは負としてそれぞれのハイブリッド評価することによって、生成する 2 進符号をアドレス (<http://www-genome.wi.mit.edu/cgi-bin/contig/rhmapper.pl>) にあるサーバを用いてフレームワーク地図を形成するマーカーについての類似コードと比較し、本遺伝子の染色体上の位置を同定した。その結果、「TCoA1」はマーカー D17S1557 に存在することが確認された(図 2B)。「TCoA1」が D17S1557 から離れた場所に位置する

可能性を示す 11 以下のスコアが唯一認められた。この場所は FISH により染色体第 17 番 q24 に FAC1 があるという結果(Bowser (1996) Genomics 38, 455-457)と一致する。

「TCoA1」の更に正確な位置を知るために、CEPH mega-YAC ライブラリーを用い、プライマー nb15S (AAGATGTTGTCTTGGAGCCGT/配列番号: 7)とプライマー nb15Q (TTTTTTACCATTGCTTCAGTCCC/配列番号: 8)を用い hierarchical PCR でスクリーニングを行った(Jones et al.(1994) Genomics 24, 266-275)。シングルクローン 983d12 が同定され、CEPH infoclone データベース (www.cephb.fr/infoclone.html)によって YAC 983d12 の地図情報を検索したが、このクローンに関する情報は全く得られなかった。しかし、Alu-PCR 産物のハイブリダイゼーションにより、983d12 と部分的に重複する 2 つのクローン (902c10 と 938f7)は、ともに D17S1557 の末端にある D17S789 に対して陽性であることが認められた。これは本発明者らにより得られた放射ハイブリッドによる結果と一致し、細胞遺伝学上「TCoA1」が染色体第 17 番 q23 に位置することを示すものであり (Collins et al. (1996) Proc. Natl. Acad. Sci. USA 93, 14771-14775)、FAC1 において報告されている染色体第 17 番 q24(Bowser (1996) Genomics 38, 455-457)に僅かながら近接した位置である。

【実施例 3】 「TCoA1」の発現解析

240 bp の cDNA プロブを用い、16 の正常組織をパネルとしてノーザン解析を行った。プロブはランダムプライミングにより [α - 32 P] dCTP で標識し、クロマスピン 10 カラム (Clontech 社)で精製した。ノーザン分析のためハイブリダイゼーションは ExpressHyb ハイブリダイゼーション溶液(Clontech 社)を用い、65°C で 1 時間かけて行った。フィルターを 65°C で 0.5xSSC、0.1% SDS の最終的ストレンジェンシーになるまで洗浄した。続いてそのフィルターをシグナル強度増大のために -70°C で 1-3 日間オートラジオグラフィーを行った。mRNA プロットは Clontech 社から購入し、300-540 塩基の領域に相当する 240bp の

cDNA プローブでハイブリダイズした。その結果、ほとんどの組織で、約 10.5kb の mRNA が検出され、この転写産物の大きさは、塩基配列から同定された ORF の大きさと同一で、報告された FAC1 の結果と一致した(図 3)。

〔実施例 4〕 TCoA1 完全長 cDNA 塩基配列の決定

本発明者等は、完全な cDNA を単離するために、実施例 1 (2) の 119bp の PCR 産物をプローブとして精巢 cDNA ライブラリー(Clontech 社 HL3024a)の再スクリーニングを行なった。スクリーニングは、実施例 1 (2) と同様の条件で行なった。

これにより得られた cDNA の塩基配列を解読した結果、開始コドンメチオニンの上流にインフレームの終止コドンが存在する 9700 塩基の配列であった。このことから、得られた cDNA が、全長配列であることが判明した。単離した完全長 cDNA の塩基配列を配列番号：9 に、該 cDNA がコードするタンパク質のアミノ酸配列を配列番号：10 に示す。

TCoA1 の塩基配列を FAC-1 (Zhu and Bowser, (1996) *Biochemica et Biophysica Acta* 1309, 5-8) と比較すると、FAC-1 の 57-1519 位の塩基配列が TCoA1 の 461-1917 位の塩基配列に、また FAC-1 の 1898-2622 位が TCoA1 の 1918-2643 位にほぼ一致しすることが判明した。しかし FAC-1 の 1520-1897 位は TCoA1 の塩基配列には存在しない。TCoA1 の塩基配列には 2781 のアミノ酸をコードするオープンリーディングフレーム (ORF) が存在するのに対して、FAC-1 の塩基配列には開始コドンメチオニンで始まる TCoA1 の一部のわずか 810 アミノ酸に相当する ORF が存在するのみである。TCoA1 のアミノ酸配列は 2 つの C4HC3 ジンクフィンガー (254-295 位のアミノ酸) と 1 つのプロモドメイン (2684-2747 位のアミノ酸) を保持している。さらに広範囲のグルタミン酸リッチな領域 (1840-2400 位のアミノ酸) がある。

〔実施例 5〕 TCoA1 の N 末端領域と相互作用するタンパク質の同定

TCoA1 の C4HC3 ジンクフィンガーを含む最初の 482 アミノ酸をコードしてい

る cDNA クローンを用いて、マウス精巣とヒト脳の酵母 2 ハイブリッド cDNA ライブラリー (Clontech 社, Palo Alto) をスクリーニングした。また、この酵母 2 ハイブリッドライブラリースクリーニングにはクローンテックの手法に従って、酵母ベクター PJ69-4A (James P, et al. (1996) *Genetics*;144(4):1425-36) を用いて行った。

その結果、ヒトの cDNA ライブラリーから hSNF2H 遺伝子 (Aihara T, et al. (1998) *Cytogenet Cell Genet*;81(3-4):191-3) が単離され、これに対応するマウス遺伝子がマウスの cDNA ライブラリーから単離された。

また、hSNF2L 遺伝子が、ヒトの cDNA ライブラリーから同定された (Aihara T, et al. (1998) *Cytogenet Cell Genet*;81(3-4):191-3)。

さらに、ヒトの cDNA ライブラリーから、転写コアクチベーターである NCoA-62 (Skip としても知られている) (Baudino TA, et al. (1998) *J Biol Chem*; 273(26):16434-41、Dahl R, et al. (1998) *Oncogene*;16(12):1579-86) が単離された。

hSNF2H/2L は *D.Melanogaster* ISWI のヒト相同体である。この ISWI タンパク質はクロマチン再構成複合体中で発見されており、この複合体は DNA 上で ATPase 依存性にヌクレオソームを再構成する分子機械であることが報告されている (Varga-Weisz PD, Becker PB. (1998) *Curr Opin Cell Biol*;10(3):346-53)。hSNF2H や hSNF2L はこれらの複合体の中で ATPase サブユニットとして働いている。

最近、ISWI だけで染色体の再構成の活性がある可能性を示唆する報告もなされている (Corona DF, et al. (1999) *Mol Cell* 3(2):239-45)。取得した hSNF2H クローンの全長配列は、データベースの配列 (GenBank Accession No.AB010882) と比較すると C 末端側 50 アミノ酸の欠失があり、オルタナティブスプライシングをしていると考えられた。

NCoA-62/Skip は、種々の核受容体 (VDR、RAR) のリガンド結合ドメインやウ

イルス腫瘍タンパク質 (oncoprotein) である Ski と相互作用する転写コアクチベーターである。またこの NCoA-62/Skip はエクジソンにより誘導されるショウジョウバエ (*Drosophila*) の Bx42 タンパク質に相同性がある。

TCoA1 と前述のタンパク質との相互作用を確認するために、クローンテックの手法に従っての哺乳動物 2 ハイブリッドシステムのコンストラクト (Clontech, Palo Alto) を使って解析した。その結果 TCoA1 と hSNF2H の間には特異的な相互作用を見つけることができたが (図 4)、hSNF2L や Skip に対してはいかなる相互作用も見られなかった。hSNF2H と hSNF2L の類似性から判断して、hSNF2L の相互作用の欠如は驚くべきことで、このシステムで hSNF2L の発現がなかったことが推測される。

【実施例 6】 TCoA1 の相互作用地図

酵母 2 ハイブリッドシステムを用い、相互作用地図作製の実験を以下のように行った。TCoA1 の種々の領域 (図 5 参照) をコードする cDNA を pAS ベクター (Clontech 社) にクローニングした。また、TCoA1 の種々の領域との相互作用の検出を行なう 3 つのタンパク質 (hSNF2H, hSNF2L および NCoA-62/Skip) をコードする cDNA を pACT ベクター (Clontech 社) にクローニングした。これらベクターの組み合わせを、酵母宿主 (PJ69-4A) に導入し、該宿主内で発現させたタンパク質同士の相互作用をルシフェラーゼをレポーターとして検出した。

その結果、hSNF2H、hSNF2L、および bx42 (NCoA-62/Skip) のいずれとも相互作用する TCoA1 の領域が明らかになった。即ち、図 5 で見られるように、これら 3 つのすべてのタンパク質は TCoA1 の 85-247 アミノ酸と相互作用した。

この事実から、さらに、タンパク質相互作用部位として知られている C4HC3 ジンクフィンガーが、これらの 3 つのタンパク質との相互作用部位から除外されていることが判明した。

【実施例 7】 プロモドメイン相互作用タンパク質の機能解析

酵母2ハイブリッドスクリーニングで同定した TCoA1 が相互作用しているクローン(hSNF2H, hSNF2L, NCoA-62/Skip)は、大きなポリペプチドをコードしている。そこで、本発明者等は、次ぎに、酵母2ハイブリッドシステムを利用して、これらタンパク質中の、TCoA1 タンパク質と相互作用する領域の同定を行った。具体的には、NCoA-62 と hSNF2H 中の部分的に重なり合う一連のポリペプチドをコードする cDNA を含む pACT ベクター(Clontech 社)を構築し(図6、図7)、TCoA1 タンパク質の 1-525 位のアミノ酸をコードする cDNA を含む pAS ベクター(Clontech 社)とともに酵母細胞(PJ69-4A)に導入し、該宿主内で発現させたタンパク質同士の相互作用をルシフェラーゼをレポーターとして検出した。

NCoA-62 についてはそのオリジナルクローンの完全なカルボキシ末端ドメインを含む約 450 アミノ酸の領域、および該領域における一連の 5 つの欠失クローンについて検討した(図5)。その結果、NCoA-62 中の 224-317 位のアミノ酸が TCoA1 への相互作用領域として同定された。

一方、hSNF2H については、3 つの欠失クローンにつき検討した(図7)。その結果、hSNF2H 中の TCoA1 タンパク質との相互作用領域は、そのカルボキシ末端側にマッピングされた(921-1017 位)。同じ領域を含むもうひとつのクローン(855-1017 位)では相互作用がみられなかった。これは、このクローンが特別な二次構造を形成することによる可能性がある。

産業上の利用の可能性

本発明の転写調節因子および該転写調節因子をコードする DNA は、癌や他の細胞増殖性疾患の治療や医薬品候補化合物のスクリーニングに用いることが可能である。また、本発明の転写調節因子に結合する抗体や該転写調節因子の機能を調節する化合物、該転写調節因子の他のタンパク質との相互作用を阻害する化合物には、これら疾患の治療薬や予防薬としての利用が考えられる。

請求の範囲

1. 配列番号：1または10に記載のアミノ酸配列からなるタンパク質。
2. 配列番号：1または10に記載のアミノ酸配列において1若しくは複数のアミノ酸が置換、欠失、付加、および／または挿入されたアミノ酸配列からなり、プロモドメインを有する転写調節因子。
3. 配列番号：1または10に記載のアミノ酸配列において1若しくは複数のアミノ酸が置換、欠失、付加、および／または挿入されたアミノ酸配列からなり、hSNF2H、hSNF2L、NCoA-62/Skip およびそれらの他種相同体からなる群より選択されるタンパク質に結合する活性を有するタンパク質。
4. 配列番号：2または9に記載の塩基配列からなる DNA とハイブリダイズする DNA がコードし、プロモドメインを有する転写調節因子。
5. 配列番号：2または9に記載の塩基配列からなる DNA とハイブリダイズする DNA がコードし、hSNF2H、hSNF2L、NCoA-62/Skip およびそれらの他種相同体からなる群より選択されるタンパク質に結合する活性を有する転写調節因子。
6. 請求項1から5のいずれかに記載の転写調節因子をコードする DNA。
7. 配列番号：2または9に記載の塩基配列のコード領域を含む、請求項6に記載の DNA。
8. 請求項6または7に記載の DNA を含むベクター。
9. 請求項6または7に記載の DNA を発現可能に保持する形質転換体。
10. 請求項9に記載の形質転換体を培養する工程を含む、請求項1から5のいずれかに記載の転写調節因子の製造方法。
11. 請求項1から5のいずれかに記載の転写調節因子に結合する抗体。
12. 請求項1から5のいずれかに記載の転写調節因子に結合する活性を有する化合物のスクリーニング方法であって、

(a) 該転写調節因子と被験サンプルとを接触させる工程、
(b) 該転写調節因子と被験サンプルとの結合活性を検出する工程、
(c) 該転写調節因子に結合する活性を有する化合物を選択する工程、を含む方法。

13. 請求項1から5のいずれかに記載の転写調節因子と、hSNF2H、hSNF2L、NCoA-62/Skip およびそれらの他種相同体からなる群より選択されるタンパク質との結合を阻害または促進する化合物をスクリーニングする方法であって、

(a) 被検サンプルの存在下で、該転写調節因子と hSNF2H、hSNF2L、NCoA-62/Skip またはそれらの他種相同体とを接触させる工程、

(b) 該転写調節因子と hSNF2H、hSNF2L、NCoA-62/Skip またはそれらの他種相同体との結合活性を検出する工程、

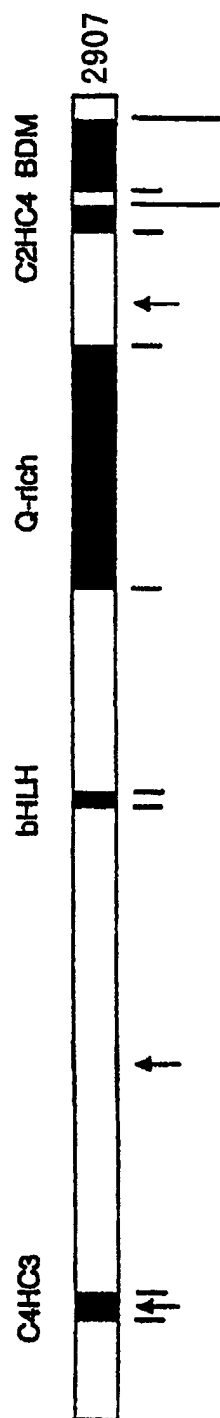
(c) 被検サンプル非存在下で検出した場合（対照）と比較して、該結合活性を低下または増加させる化合物を選択する工程、を含む方法。

14. 請求項13に記載の方法により単離しうる、請求項1から5のいずれかに記載の転写調節因子と、hSNF2H、hSNF2L、NCoA-62/Skip およびそれらの他種相同体からなる群より選択されるタンパク質との結合を阻害する化合物。

15. 配列番号：2または9に記載の塩基配列からなる DNA と特異的にハイブリダイズし、少なくとも15ヌクレオチドの鎖長を有する DNA。

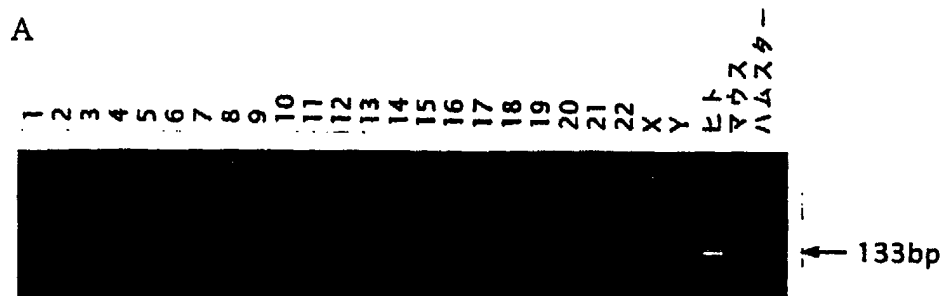
1 / 8

☒ 1

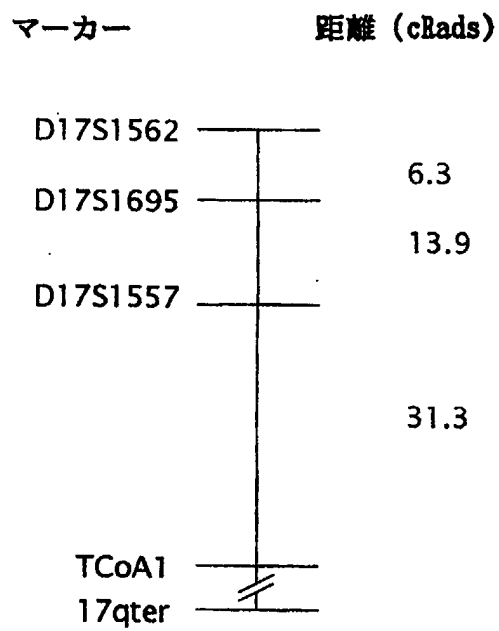


2 / 8

図 2

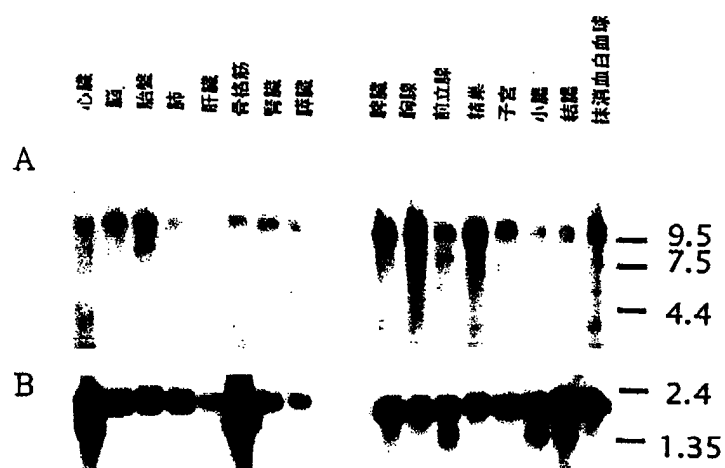


B



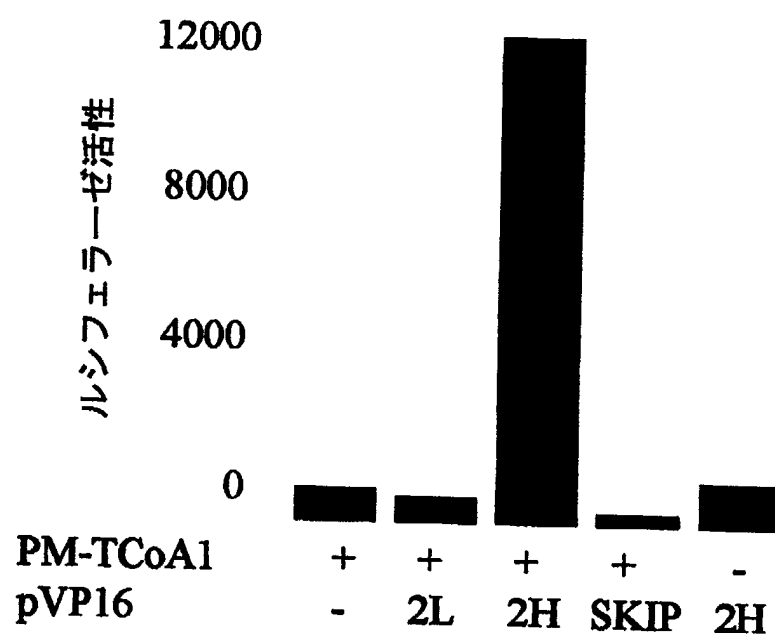
3 / 8

図 3



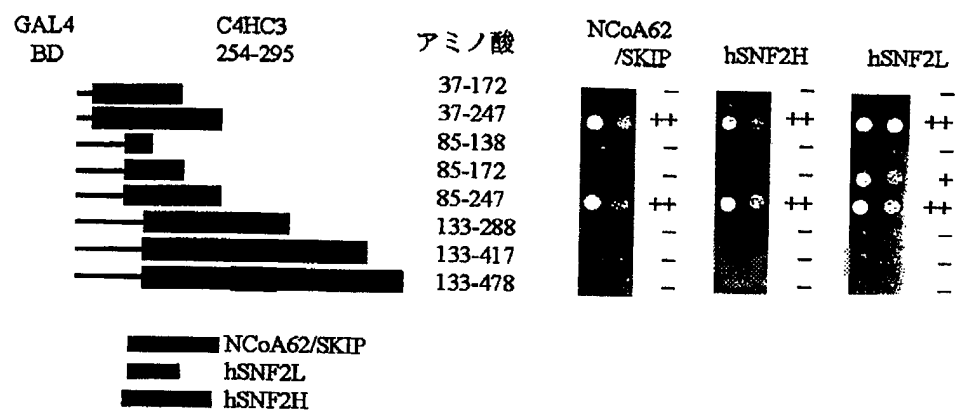
4 / 8

図 4



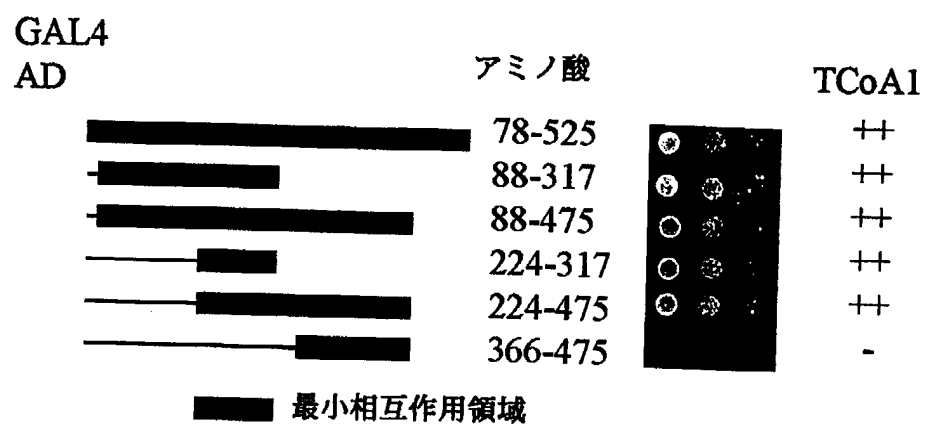
5 / 8

図 5



6 / 8

図 6









7 / 8


図 7

GAL4
AD

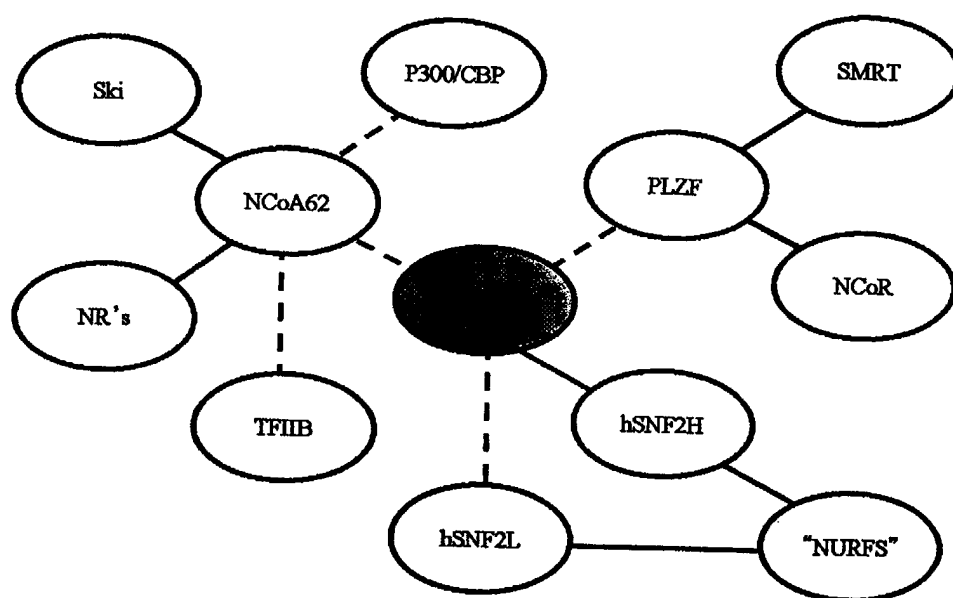
アミノ酸

TCoA1

	692-863		-
	855-1017		-
	921-1017		++

 最小相互作用領域

8



配列表

SEQUENCE LISTING

<110> CHUGAI RESEARCH INSTITUTE FOR MOLECULAR MEDICINE, INC.

株式会社中外分子医学研究所

<120> TRANSCRIPTION FACTORS

転写調節因子

<130> C2-001PCT

<140>

<141>

<150> JP 10-137631

<151> 1998-04-30

<160> 10

<170> PatentIn Ver. 2.0

<210> 1

<211> 2907

<212> PRT

<213> Homo sapiens

<400> 1

Met Val Ser Glu Glu Glu Glu Glu Glu Asp Gly Asp Ala Glu Glu Thr

1 5 10 15

Gln Asp Ser Glu Asp Asp Glu Glu Asp Glu Met Glu Glu Asp Asp Asp

20 25 30

Asp Ser Asp Tyr Pro Glu Glu Met Glu Asp Asp Asp Asp Ala Ser

35 40 45

Tyr Cys Thr Glu Ser Ser Phe Arg Ser His Ser Thr Tyr Ser Ser Thr

50 55 60

Pro Gly Arg Arg Lys Pro Arg Val His Arg Pro Arg Ser Pro Ile Leu

65 70 75 80

Glu Glu Lys Asp Ile Pro Pro Leu Glu Phe Pro Lys Ser Ser Glu Asp

85 90 95

Leu Met Val Pro Asn Glu His Ile Met Asn Val Ile Ala Ile Tyr Glu

100 105 110

Val Leu Arg Asn Phe Gly Thr Val Leu Arg Leu Ser Pro Phe Arg Phe

115 120 125

Glu Asp Phe Cys Ala Ala Leu Val Ser Gln Glu Gln Cys Thr Leu Met

130

135

140

Ala Glu Met His Val Val Leu Leu Lys Ala Val Leu Arg Glu Glu Asp

145

150

155

160

Thr Ser Asn Thr Thr Phe Gly Pro Ala Asp Leu Lys Asp Ser Val Asn

165

170

175

Ser Thr Leu Tyr Phe Ile Asp Gly Met Thr Trp Pro Glu Val Leu Arg

180

185

190

Val Tyr Cys Glu Ser Asp Lys Glu Tyr His His Val Leu Pro Tyr Gln

195

200

205

Glu Ala Glu Asp Tyr Pro Tyr Gly Pro Val Glu Asn Lys Ile Lys Val

210

215

220

Leu Gln Phe Leu Val Asp Gln Phe Leu Thr Thr Asn Ile Ala Arg Glu

225

230

235

240

Glu Leu Met Ser Glu Gly Val Ile Gln Tyr Asp Asp His Cys Arg Val

245

250

255

Cys His Lys Leu Gly Asp Leu Leu Cys Cys Glu Thr Cys Ser Ala Val

260

265

270

Tyr His Leu Glu Cys Val Lys Pro Pro Leu Glu Glu Val Pro Glu Asp

275

280

285

Glu Trp Gln Cys Glu Val Cys Val Ala His Lys Val Pro Gly Val Thr

290

295

300

Asp Cys Val Ala Glu Ile Gln Lys Asn Lys Pro Tyr Ile Arg His Glu

305

310

315

320

Pro Ile Gly Tyr Asp Arg Ser Arg Arg Lys Tyr Trp Phe Leu Asn Arg

325

330

335

Arg Leu Ile Ile Glu Glu Asp Thr Glu Asn Glu Asn Glu Lys Lys Ile

340

345

350

Trp Tyr Tyr Ser Thr Lys Val Gln Leu Ala Glu Leu Ile Asp Cys Leu

355

360

365

Asp Lys Asp Tyr Trp Glu Ala Glu Leu Cys Lys Ile Leu Glu Glu Met

370

375

380

Arg Glu Glu Ile His Arg His Met Asp Ile Thr Glu Asp Leu Thr Asn

385

390

395

400

Lys Ala Arg Gly Ser Asn Lys Ser Phe Leu Ala Ala Ala Asn Glu Glu

405	410	415	
Ile Leu Glu Ser Ile Arg Ala Lys Lys Gly Asp Ile Asp Asn Val Lys			
420	425	430	
Ser Pro Glu Glu Thr Glu Lys Asp Lys Asn Glu Thr Glu Asn Asp Ser			
435	440	445	
Lys Asp Ala Glu Lys Asn Arg Glu Glu Phe Glu Asp Gln Ser Leu Glu			
450	455	460	
Lys Asp Ser Asp Asp Lys Thr Pro Asp Asp Asp Pro Glu Gln Gly Lys			
465	470	475	480
Ser Glu Glu Pro Thr Glu Val Gly Asp Lys Gly Asn Ser Val Ser Ala			
485	490	495	
Asn Leu Gly Asp Asn Thr Thr Asn Ala Thr Ser Glu Glu Thr Ser Pro			
500	505	510	
Ser Glu Gly Arg Ser Pro Val Gly Cys Leu Ser Glu Thr Pro Asp Ser			
515	520	525	
Ser Asn Met Ala Glu Lys Lys Val Ala Ser Glu Leu Pro Gln Asp Val			
530	535	540	

Pro Glu Glu Pro Asn Lys Thr Cys Glu Ser Ser Asn Thr Ser Ala Thr
545 550 555 560

Thr Thr Ser Ile Gln Pro Asn Leu Glu Asn Ser Asn Ser Ser Ser Glu
565 570 575

Leu Asn Ser Ser Gln Ser Glu Ser Ala Lys Ala Ala Asp Asp Pro Glu
580 585 590

Asn Gly Glu Arg Glu Ser His Thr Pro Val Ser Ile Gln Glu Glu Ile
595 600 605

Val Gly Asp Phe Thr Ser Glu Lys Ser Thr Gly Glu Leu Ser Glu Ser
610 615 620

Pro Gly Ala Gly Lys Gly Ala Ser Gly Ser Thr Arg Ile Ile Thr Arg
625 630 635 640

Leu Arg Asn Pro Asp Ser Lys Leu Ser Gln Leu Lys Ser Gln Gln Val
645 650 655

Ala Ala Ala Ala His Glu Ala Asn Lys Leu Phe Lys Glu Gly Lys Glu
660 665 670

Val Leu Val Val Asn Ser Gln Gly Glu Ile Ser Arg Leu Ser Thr Lys
675 680 685

Lys Glu Val Ile Met Lys Gly Asn Ile Asn Asn Tyr Phe Lys Leu Gly

690

695

700

Gln Glu Gly Lys Tyr Arg Val Tyr His Asn Gln Tyr Ser Thr Asn Ser

705

710

715

720

Phe Ala Leu Asn Lys His Gln His Arg Glu Asp His Asp Lys Arg Arg

725

730

735

His Leu Ala His Lys Phe Cys Leu Thr Pro Ala Gly Glu Phe Lys Trp

740

745

750

Asn Gly Ser Val His Gly Ser Lys Val Leu Thr Ile Ser Thr Leu Arg

755

760

765

Leu Thr Ile Thr Gln Leu Glu Asn Asn Ile Pro Ser Ser Phe Phe His

770

775

780

Pro Asn Trp Ala Ser His Arg Ala Asn Trp Ile Lys Ala Val Gln Met

785

790

795

800

Cys Ser Lys Pro Arg Glu Phe Ala Leu Ala Leu Ala Ile Leu Glu Cys

805

810

815

Ala Val Lys Pro Val Val Met Leu Pro Ile Trp Arg Glu Phe Leu Gly

820	825	830
His Thr Arg Leu His Arg Met Thr Ser Ile Glu Arg Glu Glu Lys Glu		
835	840	845
Lys Val Lys Lys Lys Glu Lys Lys Gln Glu Glu Glu Glu Thr Met Gln		
850	855	860
Gln Ala Thr Trp Val Lys Tyr Thr Phe Pro Val Lys His Gln Val Trp		
865	870	875 880
Lys Gln Lys Gly Glu Glu Tyr Arg Val Thr Gly Tyr Gly Gly Trp Ser		
885	890	895
Trp Ile Ser Lys Thr His Val Tyr Arg Phe Val Pro Lys Leu Pro Gly		
900	905	910
Asn Thr Asn Val Asn Tyr Arg Lys Ser Leu Glu Gly Thr Lys Asn Asn		
915	920	925
Met Asp Glu Asn Met Asp Glu Ser Asp Lys Arg Lys Cys Ser Arg Ser		
930	935	940
Pro Lys Lys Ile Lys Ile Glu Pro Asp Ser Glu Lys Asp Glu Val Lys		
945	950	955 960

Gly Ser Asp Ala Ala Lys Gly Ala Asp Gln Asn Glu Met Asp Ile Ser

965

970

975

Lys Ile Thr Glu Lys Lys Asp Gln Asp Val Lys Glu Leu Leu Asp Ser

980

985

990

Asp Ser Asp Lys Pro Cys Lys Glu Glu Pro Met Glu Val Asp Asp Asp

995

1000

1005

Met Lys Thr Glu Ser His Val Asn Cys Gln Glu Ser Ser Gln Val Asp

1010

1015

1020

Val Val Asn Val Ser Glu Gly Phe His Leu Arg Thr Ser Tyr Lys Lys

025

1030

1035

1040

Lys Thr Lys Ser Ser Lys Leu Asp Gly Leu Leu Glu Arg Arg Ile Lys

1045

1050

1055

Gln Phe Thr Leu Glu Glu Lys Gln Arg Leu Glu Lys Ile Lys Leu Glu

1060

1065

1070

Gly Gly Ile Lys Gly Ile Gly Lys Thr Ser Thr Asn Ser Ser Lys Asn

1075

1080

1085

Leu Ser Glu Ser Pro Val Ile Thr Lys Ala Lys Glu Gly Cys Gln Ser

1090

1095

1100

Asp Ser Met Arg Gln Glu Gln Ser Pro Asn Ala Asn Asn Asp Gln Pro

105

1110

1115

1120

Glu Asp Leu Ile Gln Gly Cys Ser Gln Ser Asp Ser Ser Val Leu Arg

1125

1130

1135

Met Ser Asp Pro Ser His Thr Thr Asn Lys Leu Tyr Pro Lys Asp Arg

1140

1145

1150

Val Leu Asp Asp Val Ser Ile Arg Ser Pro Glu Thr Lys Cys Pro Lys

1155

1160

1165

Gln Asn Ser Ile Glu Asn Asp Ile Glu Glu Lys Val Ser Asp Leu Ala

1170

1175

1180

Ser Arg Gly Gln Glu Pro Thr Lys Ser Lys Thr Lys Gly Asn Asp Phe

185

1190

1195

1200

Phe Ile Asp Asp Ser Lys Leu Ala Ser Ala Asp Asp Ile Gly Thr Leu

1205

1210

1215

Ile Cys Lys Asn Lys Lys Pro Leu Ile Gln Glu Glu Ser Asp Thr Ile

1220

1225

1230

Val Ser Ser Ser Lys Ser Ala Leu His Ser Ser Val Pro Lys Ser Thr

1235	1240	1245
Asn Asp Arg Asp Ala Thr Pro Leu Ser Arg Ala Met Asp Phe Glu Gly		
1250	1255	1260
Lys Leu Gly Cys Asp Ser Glu Ser Asn Ser Thr Leu Glu Asn Ser Ser		
265	1270	1275 1280
Asp Thr Val Ser Ile Gln Asp Ser Ser Glu Glu Asp Met Ile Val Gln		
1285	1290	1295
Asn Ser Asn Glu Ser Ile Ser Glu Gln Phe Arg Thr Arg Glu Gln Asp		
1300	1305	1310
Val Glu Val Leu Glu Pro Leu Lys Cys Glu Leu Val Ser Gly Glu Ser		
1315	1320	1325
Thr Gly Asn Cys Glu Asp Arg Leu Pro Val Lys Gly Thr Glu Ala Asn		
1330	1335	1340
Gly Lys Lys Pro Ser Gln Gln Lys Lys Leu Glu Glu Arg Pro Val Asn		
345	1350	1355 1360
Lys Cys Ser Asp Gln Ile Lys Leu Lys Asn Thr Thr Asp Lys Lys Asn		
1365	1370	1375

Asn Glu Asn Arg Glu Ser Glu Lys Lys Gly Gln Arg Thr Ser Thr Phe

1380

1385

1390

Gln Ile Asn Gly Lys Asp Asn Lys Pro Lys Ile Tyr Leu Lys Gly Glu

1395

1400

1405

Cys Leu Lys Glu Ile Ser Glu Ser Arg Val Val Ser Gly Asn Val Glu

1410

1415

1420

Pro Lys Val Asn Asn Ile Asn Lys Ile Ile Pro Glu Asn Asp Ile Lys

425

1430

1435

1440

Ser Leu Thr Val Lys Glu Ser Ala Ile Arg Pro Phe Ile Asn Gly Asp

1445

1450

1455

Val Ile Met Glu Asp Phe Asn Glu Arg Asn Ser Ser Glu Thr Lys Ser

1460

1465

1470

His Leu Leu Ser Ser Ser Asp Ala Glu Gly Asn Tyr Arg Asp Ser Leu

1475

1480

1485

Glu Thr Leu Pro Ser Thr Lys Glu Ser Asp Ser Thr Gln Thr Thr Thr

1490

1495

1500

Pro Ser Ala Ser Cys Pro Glu Ser Asn Ser Val Asn Gln Val Glu Asp

505

1510

1515

1520

Met Glu Ile Glu Thr Ser Glu Val Lys Lys Val Thr Ser Ser Pro Ile

1525

1530

1535

Thr Ser Glu Glu Glu Ser Asn Leu Ser Asn Asp Phe Ile Asp Glu Asn

1540

1545

1550

Gly Leu Pro Ile Asn Lys Asn Glu Asn Val Asn Gly Glu Ser Lys Arg

1555

1560

1565

Lys Thr Val Ile Thr Glu Val Thr Thr Met Thr Ser Thr Val Ala Thr

1570

1575

1580

Glu Ser Lys Thr Val Ile Lys Val Glu Lys Gly Asp Lys Gln Thr Val

585

1590

1595

1600

Val Ser Ser Thr Glu Asn Cys Ala Lys Ser Thr Val Thr Thr Thr Thr

1605

1610

1615

Thr Thr Val Thr Lys Leu Ser Thr Pro Ser Thr Gly Gly Ser Val Asp

1620

1625

1630

Ile Ile Ser Val Lys Glu Gln Ser Lys Thr Val Val Thr Thr Thr Val

1635

1640

1645

Thr Asp Ser Leu Thr Thr Thr Gly Gly Thr Leu Val Thr Ser Met Thr

1650	1655	1660	
Val Ser Lys Glu Tyr Ser Thr Arg Asp Lys Val Lys Leu Met Lys Phe			
665	1670	1675	1680
Ser Arg Pro Lys Lys Thr Arg Ser Gly Thr Ala Leu Pro Ser Tyr Arg			
1685	1690	1695	
Lys Phe Val Thr Lys Ser Thr Lys Lys Ser Ile Phe Val Leu Pro Asn			
1700	1705	1710	
Asp Asp Leu Lys Lys Leu Ala Arg Lys Gly Gly Ile Arg Glu Val Pro			
1715	1720	1725	
Tyr Phe Asn Tyr Asn Ala Lys Pro Ala Leu Asp Ile Trp Pro Tyr Pro			
1730	1735	1740	
Ser Pro Arg Pro Thr Phe Gly Ile Thr Trp Arg Tyr Arg Leu Gln Thr			
745	1750	1755	1760
Val Lys Ser Leu Ala Gly Val Ser Leu Met Leu Arg Leu Leu Trp Ala			
1765	1770	1775	
Ser Leu Arg Trp Asp Asp Met Ala Ala Lys Val Pro Pro Gly Gly Gly			
1780	1785	1790	

Ser Thr Arg Thr Glu Thr Ser Glu Thr Glu Ile Thr Thr Thr Glu Ile

1795

1800

1805

Ile Lys Arg Arg Asp Val Gly Pro Tyr Gly Ile Arg Phe Glu Tyr Cys

1810

1815

1820

Ile Arg Lys Ile Ile Cys Pro Ile Gly Val Pro Glu Thr Pro Lys Glu

825

1830

1835

1840

Thr Pro Thr Pro Gln Arg Lys Gly Leu Arg Ser Ser Ala Leu Arg Pro

1845

1850

1855

Lys Arg Pro Glu Thr Pro Lys Gln Thr Gly Pro Val Ile Ile Glu Thr

1860

1865

1870

Trp Val Ala Glu Glu Glu Leu Glu Leu Trp Glu Ile Arg Ala Phe Ala

1875

1880

1885

Glu Arg Val Glu Lys Glu Lys Ala Gln Ala Val Glu Gln Gln Ala Lys

1890

1895

1900

Lys Arg Leu Glu Gln Gln Lys Pro Thr Val Ile Ala Thr Ser Thr Thr

905

1910

1915

1920

Ser Pro Thr Ser Ser Thr Thr Ser Thr Ile Ser Pro Ala Gln Lys Val

1925

1930

1935

Met Val Ala Pro Ile Ser Gly Ser Val Thr Thr Gly Thr Lys Met Val

1940

1945

1950

Leu Thr Thr Lys Val Gly Ser Pro Ala Thr Val Thr Phe Gln Gln Asn

1955

1960

1965

Lys Asn Phe His Gln Thr Phe Ala Thr Trp Val Lys Gln Gly Gln Ser

1970

1975

1980

Asn Ser Gly Val Val Gln Val Gln Gln Lys Val Leu Gly Ile Ile Pro

985

1990

1995

2000

Ser Ser Thr Gly Thr Ser Gln Gln Thr Phe Thr Ser Phe Gln Pro Arg

2005

2010

2015

Thr Ala Thr Val Thr Ile Arg Pro Asn Thr Ser Gly Ser Gly Gly Thr

2020

2025

2030

Thr Ser Asn Ser Gln Val Ile Thr Gly Pro Gln Ile Arg Pro Gly Met

2035

2040

2045

Thr Val Ile Arg Thr Pro Leu Gln Gln Ser Thr Leu Gly Lys Ala Ile

2050

2055

2060

Ile Arg Thr Pro Val Met Val Gln Pro Gly Ala Pro Gln Gln Val Met

065 2070 2075 2080

Thr Gln Ile Ile Arg Gly Gln Pro Val Ser Thr Ala Val Ser Ala Pro

 2085 2090 2095

Asn Thr Val Ser Ser Thr Pro Gly Gln Lys Ser Leu Thr Ser Ala Thr

 2100 2105 2110

Ser Thr Ser Asn Ile Gln Ser Ser Ala Ser Gln Pro Pro Arg Pro Gln

 2115 2120 2125

Gln Gly Gln Val Lys Leu Thr Met Ala Gln Leu Thr Gln Leu Thr Gln

 2130 2135 2140

Gly His Gly Gly Asn Gln Gly Leu Thr Val Val Ile Gln Gly Gln Gly

145 2150 2155 2160

Gln Thr Thr Gly Gln Leu Gln Leu Ile Pro Gln Gly Val Thr Val Leu

 2165 2170 2175

Pro Gly Pro Gly Gln Gln Leu Met Gln Ala Ala Met Pro Asn Gly Thr

 2180 2185 2190

Val Gln Arg Phe Leu Phe Thr Pro Leu Ala Thr Thr Ala Thr Thr Ala

 2195 2200 2205

Ser Thr Thr Thr Thr Thr Val Ser Thr Thr Ala Ala Gly Thr Gly Glu

2210

2215

2220

Gln Arg Gln Ser Lys Leu Ser Pro Gln Met Gln Val His Gln Asp Lys

225

2230

2235

2240

Thr Leu Pro Pro Ala Gln Ser Ser Ser Val Gly Pro Ala Lys Ala Gln

2245

2250

2255

Pro Gln Thr Ala Gln Pro Ser Ala Arg Pro Gln Pro Gln Thr Gln Pro

2260

2265

2270

Gln Ser Pro Ala Gln Pro Glu Val Gln Thr Gln Pro Glu Val Gln Thr

2275

2280

2285

Gln Thr Thr Val Ser Ser His Val Pro Ser Glu Ala Gln Pro Thr His

2290

2295

2300

Ala Gln Ser Ser Lys Pro Gln Val Ala Ala Gln Ser Gln Pro Gln Ser

305

2310

2315

2320

Asn Val Gln Gly Gln Ser Pro Val Arg Val Gln Ser Pro Ser Gln Thr

2325

2330

2335

Arg Ile Arg Pro Ser Thr Pro Ser Gln Leu Ser Pro Gly Gln Gln Ser

2340

2345

2350

Gln Val Gln Thr Thr Thr Ser Gln Pro Ile Pro Ile Gln Pro His Thr

2355

2360

2365

Ser Leu Gln Ile Pro Ser Gln Gly Gln Pro Gln Ser Gln Pro Gln Val

2370

2375

2380

Gln Ser Ser Thr Gln Thr Leu Ser Ser Gly Gln Thr Leu Asn Gln Val

385

2390

2395

2400

Ser Val Ser Ser Pro Ser Arg Pro Gln Leu Gln Ile Gln Gln Pro Gln

2405

2410

2415

Pro Gln Val Ile Ala Val Pro Gln Leu Gln Gln Gln Val Gln Val Leu

2420

2425

2430

Ser Gln Ile Gln Ser Gln Val Val Ala Gln Ile Gln Ala Gln Gln Ser

2435

2440

2445

Gly Val Pro Gln Gln Ile Lys Leu Gln Leu Pro Ile Gln Ile Gln Gln

2450

2455

2460

Ser Ser Ala Val Gln Thr His Gln Ile Gln Asn Val Val Thr Val Gln

465

2470

2475

2480

Ala Ala Ser Val Gln Glu Gln Leu Gln Arg Val Gln Gln Leu Arg Asp

2485	2490	2495	
Gln Gln Gln Lys Lys Lys Gln Gln Gln Ile Glu Ile Asn Val Asn Thr			
2500	2505	2510	
Pro Ser Lys Leu Leu Ile Lys Val Glu Ile Ile Gln Lys Gln Val Val			
2515	2520	2525	
Met Lys His Asn Ala Val Ile Glu His Leu Lys Gln Lys Lys Ser Met			
2530	2535	2540	
Thr Pro Ala Glu Arg Glu Glu Asn Gln Arg Met Ile Val Cys Asn Gln			
545	2550	2555	2560
Val Met Lys Tyr Ile Leu Asp Lys Ile Asp Lys Glu Glu Lys Gln Ala			
2565	2570	2575	
Ala Lys Lys Arg Lys Arg Glu Glu Ser Val Glu Gln Lys Arg Ser Lys			
2580	2585	2590	
Gln Asn Ala Thr Lys Leu Ser Ala Leu Leu Phe Lys His Lys Glu Gln			
2595	2600	2605	
Leu Arg Ala Glu Ile Leu Lys Lys Arg Ala Leu Leu Asp Lys Asp Leu			
2610	2615	2620	

Gln Ile Glu Val Gln Glu Glu Leu Lys Arg Asp Leu Lys Ile Lys Lys

625 2630 2635 2640

Glu Lys Asp Leu Met Gln Leu Ala Gln Ala Thr Ala Val Ala Ala Pro

2645 2650 2655

Cys Pro Pro Val Thr Pro Val Leu Pro Ala Pro Pro Ala Pro Pro Pro

2660 2665 2670

Ser Pro Pro Pro Pro Pro Gly Val Gln His Thr Gly Leu Leu Ser Thr

2675 2680 2685

Pro Thr Leu Pro Val Ala Ser Gln Lys Arg Lys Arg Glu Glu Glu Lys

2690 2695 2700

Asp Ser Ser Ser Lys Ser Lys Lys Lys Lys Met Ile Ser Thr Thr Ser

705 2710 2715 2720

Lys Glu Thr Lys Lys Asp Thr Lys Leu Tyr Cys Ile Cys Lys Thr Pro

2725 2730 2735

Tyr Asp Glu Ser Lys Phe Tyr Ile Gly Cys Asp Arg Cys Gln Asn Trp

2740 2745 2750

Tyr His Gly Arg Cys Val Gly Ile Leu Gln Ser Glu Ala Glu Leu Ile

2755 2760 2765

Asp Glu Tyr Val Cys Pro Gln Cys Gln Ser Thr Glu Asp Ala Met Thr

2770

2775

2780

Val Leu Thr Pro Leu Thr Glu Lys Asp Tyr Glu Gly Leu Lys Arg Val

785

2790

2795

2800

Leu Arg Ser Leu Gln Ala His Lys Met Ala Trp Pro Phe Leu Glu Pro

2805

2810

2815

Val Asp Pro Asn Asp Ala Pro Asp Tyr Tyr Gly Val Ile Lys Glu Pro

2820

2825

2830

Met Asp Leu Ala Thr Met Glu Glu Arg Val Gln Arg Arg Tyr Tyr Glu

2835

2840

2845

Lys Leu Thr Glu Phe Val Ala Asp Met Thr Lys Ile Phe Asp Asn Cys

2850

2855

2860

Arg Tyr Tyr Asn Pro Ser Asp Ser Pro Phe Tyr Gln Cys Ala Glu Val

865

2870

2875

2880

Leu Glu Ser Phe Phe Val Gln Lys Leu Lys Gly Phe Lys Ala Ser Arg

2885

2890

2895

Ser His Asn Asn Lys Leu Gln Ser Thr Ala Ser

2900

2905

<210> 2

<211> 9865

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (259)..(8979)

<400> 2

ggccaggctg aggtggcgcc caagacgcgg ctgagctcgc ccaggggggg cagcagtagc 60

cggaggaagc cgccgccgcc gccgccggcc cccccagca ccagcgcccc gggccggggg 120

gggcgaggag gcgggggcgg cagcagggg ggcgggggcg gcggcgccca cctgtcccgg 180

accaccgcgg ccggagggc cgtcaacaaa gtggtgtacg atgaccacga gagcgaggcg 240

gtggaggaag aggaggac atg gtc tcc gag gag gag gag gag gag gac ggc 291

Met Val Ser Glu Glu Glu Glu Glu Glu Asp Gly

1

5

10

gac gcc gag gag acc cag gat tct gag gac gac gag gag gat gag atg 339

Asp Ala Glu Glu Thr Gln Asp Ser Glu Asp Asp Glu Glu Asp Glu Met

15	20	25	
gaa gag gac gac gat gac tcc gat tat ccg gag gag atg gaa gac gac 387			
Glu Glu Asp Asp Asp Asp Ser Asp Tyr Pro Glu Glu Met Glu Asp Asp			
30	35	40	
gac gac gac gcc agt tac tgc acg gaa agc agc ttc agg agc cat agt 435			
Asp Asp Asp Ala Ser Tyr Cys Thr Glu Ser Ser Phe Arg Ser His Ser			
45	50	55	
acc tac agc agc act cca ggt agg cga aaa cca aga gta cat cgg cct 483			
Thr Tyr Ser Ser Thr Pro Gly Arg Arg Lys Pro Arg Val His Arg Pro			
60	65	70	75
cgt tct cct ata ttg gaa gaa aaa gac atc ccg ccc ctt gaa ttt ccc 531			
Arg Ser Pro Ile Leu Glu Glu Lys Asp Ile Pro Pro Leu Glu Phe Pro			
80	85	90	
aag tcc tct gag gat tta atg gtg cct aat gag cat ata atg aat gtc 579			
Lys Ser Ser Glu Asp Leu Met Val Pro Asn Glu His Ile Met Asn Val			
95	100	105	
att gcc att tac gag gta ctg cgg aac ttt ggc act gtt ttg aga tta 627			
Ile Ala Ile Tyr Glu Val Leu Arg Asn Phe Gly Thr Val Leu Arg Leu			
110	115	120	

tct cct ttt cgc ttt gag gac ttt tgt gca gct ctg gtg agc caa gag 675

Ser Pro Phe Arg Phe Glu Asp Phe Cys Ala Ala Leu Val Ser Gln Glu

125

130

135

cag tgc aca ctc atg gca gag atg cat gtt gtg ctt ttg aaa gca gtt 723

Gln Cys Thr Leu Met Ala Glu Met His Val Val Leu Leu Lys Ala Val

140

145

150

155

ctg cgt gaa gaa gac act tcc aat act acc ttt gga cct gct gat ctg 771

Leu Arg Glu Glu Asp Thr Ser Asn Thr Thr Phe Gly Pro Ala Asp Leu

160

165

170

aaa gat agc gtt aat tcc aca ctg tat ttc ata gat ggg atg acg tgg 819

Lys Asp Ser Val Asn Ser Thr Leu Tyr Phe Ile Asp Gly Met Thr Trp

175

180

185

cca gag gtg ctg cgg gtg tac tgt gag agt gat aag gag tac cat cac 867

Pro Glu Val Leu Arg Val Tyr Cys Glu Ser Asp Lys Glu Tyr His His

190

195

200

gtt ctt cct tac caa gag gca gag gac tac cca tat gga cca gta gag 915

Val Leu Pro Tyr Gln Glu Ala Glu Asp Tyr Pro Tyr Gly Pro Val Glu

205

210

215

aac aag atc aaa gtt cta cag ttt cta gtc gat cag ttt ctt aca aca 963

Asn Lys Ile Lys Val Leu Gln Phe Leu Val Asp Gln Phe Leu Thr Thr

220 225 230 235

aat att gct cga gag gaa ttg atg tct gaa ggg gtg ata cag tat gat 1011
Asn Ile Ala Arg Glu Glu Leu Met Ser Glu Gly Val Ile Gln Tyr Asp
 240 245 250

gac cat tgt agg gtt tgt cac aaa ctt ggg gat ttg ctt tgc tgt gag 1059
Asp His Cys Arg Val Cys His Lys Leu Gly Asp Leu Leu Cys Cys Glu
 255 260 265

aca tgt tca gca gta tac cat ttg gaa tgt gtg aag cca cct ctt gag 1107
Thr Cys Ser Ala Val Tyr His Leu Glu Cys Val Lys Pro Pro Leu Glu
 270 275 280

gag gtg cca gag gac gag tgg cag tgt gaa gtc tgt gta gca cac aag 1155
Glu Val Pro Glu Asp Glu Trp Gln Cys Glu Val Cys Val Ala His Lys
 285 290 295

gtg cct ggt gtg act gac tgt gtt gct gaa atc caa aaa aat aaa cca 1203
Val Pro Gly Val Thr Asp Cys Val Ala Glu Ile Gln Lys Asn Lys Pro
300 305 310 315

tat att cga cat gaa cct att gga tat gat aga agt cgg agg aaa tac 1251
Tyr Ile Arg His Glu Pro Ile Gly Tyr Asp Arg Ser Arg Arg Lys Tyr
 320 325 330

tgg ttc ttg aac cga aga ctc ata ata gaa gaa gat aca gaa aat gaa 1299

Trp Phe Leu Asn Arg Arg Leu Ile Ile Glu Glu Asp Thr Glu Asn Glu

335

340

345

aat gaa aag aaa att tgg tat tac agc aca aag gtc caa ctt gca gaa 1347

Asn Glu Lys Lys Ile Trp Tyr Tyr Ser Thr Lys Val Gln Leu Ala Glu

350

355

360

tta att gac tgt cta gac aaa gat tat tgg gaa gca gaa ctc tgc aaa 1395

Leu Ile Asp Cys Leu Asp Lys Asp Tyr Trp Glu Ala Glu Leu Cys Lys

365

370

375

att cta gaa gaa atg cgt gaa gaa atc cac cga cac atg gac ata act 1443

Ile Leu Glu Glu Met Arg Glu Glu Ile His Arg His Met Asp Ile Thr

380

385

390

395

gaa gac ctg acc aat aag gct cgg ggc agt aac aaa tcc ttt ctg gcg 1491

Glu Asp Leu Thr Asn Lys Ala Arg Gly Ser Asn Lys Ser Phe Leu Ala

400

405

410

gca gct aat gaa gaa att ttg gaa tcc ata aga gcc aaa aag gga gac 1539

Ala Ala Asn Glu Glu Ile Leu Glu Ser Ile Arg Ala Lys Lys Gly Asp

415

420

425

att gat aat gtt aaa agc cca gaa gaa aca gaa aaa gac aag aat gag 1587

Ile Asp Asn Val Lys Ser Pro Glu Glu Thr Glu Lys Asp Lys Asn Glu

430 435 440

act gag aat gac tct aaa gat gct gag aaa aac aga gaa gaa ttt gaa 1635
Thr Glu Asn Asp Ser Lys Asp Ala Glu Lys Asn Arg Glu Glu Phe Glu

445 450 455

gac cag tcc ctt gaa aaa gac agt gac gac aaa aca cca gat gat gac 1683
Asp Gln Ser Leu Glu Lys Asp Ser Asp Asp Lys Thr Pro Asp Asp Asp

460 465 470 475

cct gag caa gga aaa tct gag gag cca aca gaa gtt ggg gat aaa ggt 1731
Pro Glu Gln Gly Lys Ser Glu Glu Pro Thr Glu Val Gly Asp Lys Gly

480 485 490

aac tct gtg tca gca aat ctt ggc gac aac aca aca aat gca act tca 1779
Asn Ser Val Ser Ala Asn Leu Gly Asp Asn Thr Thr Asn Ala Thr Ser

495 500 505

gaa gag act agt ccc tct gaa ggg agg agc cct gtg ggg tgt etc tca 1827
Glu Glu Thr Ser Pro Ser Glu Gly Arg Ser Pro Val Gly Cys Leu Ser

510 515 520

gaa acc ccc gat agc agc aac atg gca gag aag aag gtg gca tct gag 1875
Glu Thr Pro Asp Ser Ser Asn Met Ala Glu Lys Lys Val Ala Ser Glu

525 530 535

ctc ccc cag gat gtg cca gaa gaa cct aac aag aca tgt gag agc agt 1923

Leu Pro Gln Asp Val Pro Glu Glu Pro Asn Lys Thr Cys Glu Ser Ser

540

545

550

555

aac act agt gct acc act acc tcc atc cag cct aat ctg gaa aac agt 1971

Asn Thr Ser Ala Thr Thr Thr Ser Ile Gln Pro Asn Leu Glu Asn Ser

560

565

570

aac agc agc agt gaa cta aat tct tcc cag agt gaa tct gct aag gca 2019

Asn Ser Ser Ser Glu Leu Asn Ser Ser Gln Ser Glu Ser Ala Lys Ala

575

580

585

gct gat gat cct gaa aat gga gaa aga gaa tct cat aca cct gtc tct 2067

Ala Asp Asp Pro Glu Asn Gly Glu Arg Glu Ser His Thr Pro Val Ser

590

595

600

att cag gaa gag ata gta ggt gat ttc aca tcg gag aag tcc acc ggg 2115

Ile Gln Glu Glu Ile Val Gly Asp Phe Thr Ser Glu Lys Ser Thr Gly

605

610

615

gag cta agt gaa tct cct gga gct gga aaa gga gca tct ggc tca act 2163

Glu Leu Ser Glu Ser Pro Gly Ala Gly Lys Gly Ala Ser Gly Ser Thr

620

625

630

635

cga atc atc acc aga ttg cgg aat cca gat agc aaa ctt agt cag ctg 2211

Arg Ile Ile Thr Arg Leu Arg Asn Pro Asp Ser Lys Leu Ser Gln Leu

640	645	650	
aag agc cag cag gtg gca gcc gct gca cat gaa gca aat aaa tta ttt	2259		
Lys Ser Gln Gln Val Ala Ala Ala Ala His Glu Ala Asn Lys Leu Phe			
655	660	665	
aag gag ggc aaa gag gta ctg gta gtt aac tct caa gga gaa att tca	2307		
Lys Glu Gly Lys Glu Val Leu Val Val Asn Ser Gln Gly Glu Ile Ser			
670	675	680	
cgg ttg agc acc aaa aag gaa gtg atc atg aaa gga aat atc aac aat	2355		
Arg Leu Ser Thr Lys Lys Glu Val Ile Met Lys Gly Asn Ile Asn Asn			
685	690	695	
tat ttt aaa ttg ggt caa gaa ggg aag tat cgc gtc tac cac aat caa	2403		
Tyr Phe Lys Leu Gly Gln Glu Gly Lys Tyr Arg Val Tyr His Asn Gln			
700	705	710	715
tac tcc acc aat tca ttt gct ttg aat aag cac cag cac aga gaa gac	2451		
Tyr Ser Thr Asn Ser Phe Ala Leu Asn Lys His Gln His Arg Glu Asp			
720	725	730	
cat gat aag aga agg cat ctt gca cat aag ttc tgt ctg act cca gca	2499		
His Asp Lys Arg Arg His Leu Ala His Lys Phe Cys Leu Thr Pro Ala			
735	740	745	

gga gag ttc aaa tgg aac ggt tct gtc cat ggg tcc aaa gtt ctt acc 2547

Gly Glu Phe Lys Trp Asn Gly Ser Val His Gly Ser Lys Val Leu Thr

750

755

760

ata tct act ctg aga ctg act atc acc caa tta gaa aac aac atc cct 2595

Ile Ser Thr Leu Arg Leu Thr Ile Thr Gln Leu Glu Asn Asn Ile Pro

765

770

775

tca tcc ttt ttt cat ccc aac tgg gca tca cat agg gca aat tgg atc 2643

Ser Ser Phe Phe His Pro Asn Trp Ala Ser His Arg Ala Asn Trp Ile

780

785

790

795

aag gca gtt cag atg tgt agc aaa ccc aga gaa ttt gca ttg gct tta 2691

Lys Ala Val Gln Met Cys Ser Lys Pro Arg Glu Phe Ala Leu Ala Leu

800

805

810

gcc att ttg gag tgt gca gtt aaa cca gtt gtg atg cta cca ata tgg 2739

Ala Ile Leu Glu Cys Ala Val Lys Pro Val Val Met Leu Pro Ile Trp

815

820

825

cga gaa ttt tta gga cat acc agg tta cac cgg atg aca tca att gaa 2787

Arg Glu Phe Leu Gly His Thr Arg Leu His Arg Met Thr Ser Ile Glu

830

835

840

aga gaa gaa aag gag aaa gtc aaa aaa gag aag aaa cag gaa gaa 2835

Arg Glu Glu Lys Glu Lys Val Lys Lys Lys Glu Lys Lys Gln Glu Glu

845	850	855	
gaa gaa acg atg cag caa gcg aca tgg gta aaa tac aca ttt cca gtt 2883			
Glu Glu Thr Met Gln Gln Ala Thr Trp Val Lys Tyr Thr Phe Pro Val			
860	865	870	875
aag cat cag gtt tgg aaa caa aaa ggt gaa gag tac aga gtg aca gga 2931			
Lys His Gln Val Trp Lys Gln Lys Gly Glu Glu Tyr Arg Val Thr Gly			
	880	885	890
tat ggt ggt tgg agc tgg att agt aaa act cat gtt tat agg ttt gtt 2979			
Tyr Gly Gly Trp Ser Trp Ile Ser Lys Thr His Val Tyr Arg Phe Val			
	895	900	905
cct aaa ttg cca ggc aat act aat gtg aat tac aga aag tcg tta gaa 3027			
Pro Lys Leu Pro Gly Asn Thr Asn Val Asn Tyr Arg Lys Ser Leu Glu			
	910	915	920
gga acc aaa aat aat atg gat gaa aat atg gat gag tca gat aaa aga 3075			
Gly Thr Lys Asn Asn Met Asp Glu Asn Met Asp Glu Ser Asp Lys Arg			
	925	930	935
aaa tgt tca cga agt cca aaa aaa ata aaa ata gag cct gat tct gaa 3123			
Lys Cys Ser Arg Ser Pro Lys Lys Ile Lys Ile Glu Pro Asp Ser Glu			
940	945	950	955

aaa gat gag gta aaa ggt tca gat gct gca aaa gga gca gac caa aat 3171

Lys Asp Glu Val Lys Gly Ser Asp Ala Ala Lys Gly Ala Asp Gln Asn

960

965

970

gaa atg gat atc tca aag att act gag aag aag gac caa gat gtg aag 3219

Glu Met Asp Ile Ser Lys Ile Thr Glu Lys Lys Asp Gln Asp Val Lys

975

980

985

gag ctc tta gat tct gac agt gat aaa ccc tgc aag gaa gaa cca atg 3267

Glu Leu Leu Asp Ser Asp Ser Asp Lys Pro Cys Lys Glu Glu Pro Met

990

995

1000

gaa gta gac gat gac atg aaa aca gag tca cat gta aat tgt cag gag 3315

Glu Val Asp Asp Asp Met Lys Thr Glu Ser His Val Asn Cys Gln Glu

1005

1010

1015

agt tct caa gta gat gtg gtc aat gtt agt gag ggt ttt cat cta agg 3363

Ser Ser Gln Val Asp Val Val Asn Val Ser Glu Gly Phe His Leu Arg

1020

1025

1030

1035

act agt tac aaa aag aaa aca aaa tca tcc aaa cta gat gga ctt ctt 3411

Thr Ser Tyr Lys Lys Lys Thr Lys Ser Ser Lys Leu Asp Gly Leu Leu

1040

1045

1050

gaa agg aga att aaa cag ttt aca ctg gaa gaa aaa cag cga ctc gaa 3459

Glu Arg Arg Ile Lys Gln Phe Thr Leu Glu Glu Lys Gln Arg Leu Glu

1055	1060	1065	
aaa atc aag ttg gag ggt gga att aag ggt ata gga aag act tct aca 3507			
Lys Ile Lys Leu Glu Gly Gly Ile Lys Gly Ile Gly Lys Thr Ser Thr			
1070	1075	1080	
aat tct tca aaa aat ctc tct gaa tca cca gta ata acg aaa gca aaa 3555			
Asn Ser Ser Lys Asn Leu Ser Glu Ser Pro Val Ile Thr Lys Ala Lys			
1085	1090	1095	
gaa ggg tgt cag agt gac tcg atg aga caa gaa cag agc cca aat gca 3603			
Glu Gly Cys Gln Ser Asp Ser Met Arg Gln Glu Gln Ser Pro Asn Ala			
1100	1105	1110	1115
aat aat gat caa cct gag gac ttg att cag gga tgt tca caa agt gat 3651			
Asn Asn Asp Gln Pro Glu Asp Leu Ile Gln Gly Cys Ser Gln Ser Asp			
1120	1125	1130	
tcc tca gtt ctt aga atg agt gat cct agt cat acc aca aac aaa ctt 3699			
Ser Ser Val Leu Arg Met Ser Asp Pro Ser His Thr Thr Asn Lys Leu			
1135	1140	1145	
tat cca aaa gat cga gtg tta gat gat gtc tcc att cgg agc cca gaa 3747			
Tyr Pro Lys Asp Arg Val Leu Asp Asp Val Ser Ile Arg Ser Pro Glu			
1150	1155	1160	

aca aaa tgt ccg aaa caa aat tcc att gaa aat gac ata gaa gaa aaa 3795

Thr Lys Cys Pro Lys Gln Asn Ser Ile Glu Asn Asp Ile Glu Glu Lys

1165

1170

1175

gtc tct gac ctt gcc agt aga ggc cag gaa ccc act aag agt aaa acc 3843

Val Ser Asp Leu Ala Ser Arg Gly Gln Glu Pro Thr Lys Ser Lys Thr

1180

1185

1190

1195

aaa gga aat gat ttt ttc atc gat gac tct aaa cta gcc agt gca gat 3891

Lys Gly Asn Asp Phe Phe Ile Asp Asp Ser Lys Leu Ala Ser Ala Asp

1200

1205

1210

gat att ggt act ttg atc tgt aag aac aaa aaa ccg ctc ata cag gag 3939

Asp Ile Gly Thr Leu Ile Cys Lys Asn Lys Lys Pro Leu Ile Gln Glu

1215

1220

1225

gaa agt gac acc att gtt tct tct tcc aag agt gct tta cat tca tca 3987

Glu Ser Asp Thr Ile Val Ser Ser Ser Lys Ser Ala Leu His Ser Ser

1230

1235

1240

gtg cct aaa agt acc aat gac aga gat gcc aca cct ctg tca aga gca 4035

Val Pro Lys Ser Thr Asn Asp Arg Asp Ala Thr Pro Leu Ser Arg Ala

1245

1250

1255

atg gac ttt gaa gga aaa ctg gga tgt gac tct gaa tct aat agc act 4083

Met Asp Phe Glu Gly Lys Leu Gly Cys Asp Ser Glu Ser Asn Ser Thr

1260	1265	1270	1275
ttg gaa aat agt tct gat acc gtg tct att cag gat agc agt gaa gaa 4131			
Leu Glu Asn Ser Ser Asp Thr Val Ser Ile Gln Asp Ser Ser Glu Glu			
1280	1285	1290	
gat atg att gtt cag aat agc aat gaa agc att tct gaa cag ttc aga 4179			
Asp Met Ile Val Gln Asn Ser Asn Glu Ser Ile Ser Glu Gln Phe Arg			
1295	1300	1305	
act cga gaa caa gat gtt gaa gtc ttg gag ccg tta aag tgt gag ttg 4227			
Thr Arg Glu Gln Asp Val Glu Val Leu Glu Pro Leu Lys Cys Glu Leu			
1310	1315	1320	
gtt tct ggt gag tcc act gga aac tgt gag gac agg ctg ccg gtc aag 4275			
Val Ser Gly Glu Ser Thr Gly Asn Cys Glu Asp Arg Leu Pro Val Lys			
1325	1330	1335	
ggg act gaa gca aat ggt aaa aaa cca agt cag cag aag aaa tta gag 4323			
Gly Thr Glu Ala Asn Gly Lys Lys Pro Ser Gln Gln Lys Lys Leu Glu			
1340	1345	1350	1355
gag aga cca gtt aat aaa tgt agt gat caa ata aag cta aaa aat acc 4371			
Glu Arg Pro Val Asn Lys Cys Ser Asp Gln Ile Lys Leu Lys Asn Thr			
1360	1365	1370	

act gac aaa aag aat aat gaa aat cga gag tct gaa aag aaa gga cag 4419

Thr Asp Lys Lys Asn Asn Glu Asn Arg Glu Ser Glu Lys Lys Gly Gln

1375

1380

1385

aga aca agt aca ttt caa ata aat gga aaa gat aat aaa ccc aaa ata 4467

Arg Thr Ser Thr Phe Gln Ile Asn Gly Lys Asp Asn Lys Pro Lys Ile

1390

1395

1400

tat ttg aaa ggt gaa tgc ttg aaa gaa att tct gag agt aga gta gta 4515

Tyr Leu Lys Gly Glu Cys Leu Lys Glu Ile Ser Glu Ser Arg Val Val

1405

1410

1415

agt ggt aat gtt gaa cca aag gtt aat aat ata aat aaa ata atc cct 4563

Ser Gly Asn Val Glu Pro Lys Val Asn Asn Ile Asn Lys Ile Ile Pro

1420

1425

1430

1435

gag aat gat att aaa tca ttg act gtt aaa gaa tct gct ata agg cca 4611

Glu Asn Asp Ile Lys Ser Leu Thr Val Lys Glu Ser Ala Ile Arg Pro

1440

1445

1450

ttc att aat ggt gat gtc atc atg gaa gat ttt aat gaa aga aac agc 4659

Phe Ile Asn Gly Asp Val Ile Met Glu Asp Phe Asn Glu Arg Asn Ser

1455

1460

1465

tcc gaa aca aaa tcg cat ttg ctg agt tct tca gat gct gaa ggt aac 4707

Ser Glu Thr Lys Ser His Leu Leu Ser Ser Ser Asp Ala Glu Gly Asn

1470	1475	1480	
tac cga gat agc ctt gag acc ctg cca tca acc aaa gag tct gac agt 4755			
Tyr Arg Asp Ser Leu Glu Thr Leu Pro Ser Thr Lys Glu Ser Asp Ser			
1485	1490	1495	
aca cag acg acc aca ccc tca gca tct tgt cca gaa agc aat tca gtt 4803			
Thr Gln Thr Thr Thr Pro Ser Ala Ser Cys Pro Glu Ser Asn Ser Val			
1500	1505	1510	1515
aat cag gta gaa gat atg gaa ata gaa acc tca gaa gtt aag aaa gtt 4851			
Asn Gln Val Glu Asp Met Glu Ile Glu Thr Ser Glu Val Lys Lys Val			
1520	1525	1530	
act tca tca cct att act tct gaa gag gaa tct aat ctc agt aat gac 4899			
Thr Ser Ser Pro Ile Thr Ser Glu Glu Glu Ser Asn Leu Ser Asn Asp			
1535	1540	1545	
ttt att gat gaa aat ggt ctg ccc atc aac aaa aat gaa aat gtc aat 4947			
Phe Ile Asp Glu Asn Gly Leu Pro Ile Asn Lys Asn Glu Asn Val Asn			
1550	1555	1560	
gga gaa tct aaa aga aaa acc gtc atc aca gaa gtc acc acg atg acc 4995			
Gly Glu Ser Lys Arg Lys Thr Val Ile Thr Glu Val Thr Thr Met Thr			
1565	1570	1575	

tcc aca gtg gcc aca gaa tca aaa act gtg atc aag gta gaa aaa ggc 5043
Ser Thr Val Ala Thr Glu Ser Lys Thr Val Ile Lys Val Glu Lys Gly
1580 1585 1590 1595

gat aag caa act gtg gtt tct tcc aca gaa aat tgt gca aaa tcc act 5091
Asp Lys Gln Thr Val Val Ser Ser Thr Glu Asn Cys Ala Lys Ser Thr
1600 1605 1610

gtc aca acc acc act aca aca gtg acc aag ctt tcc aca ccc tcc aca 5139
Val Thr Thr Thr Thr Thr Thr Val Thr Lys Leu Ser Thr Pro Ser Thr
1615 1620 1625

ggc ggc agt gtg gac atc atc tct gta aag gag cag agc aaa acc gtg 5187
Gly Gly Ser Val Asp Ile Ile Ser Val Lys Glu Gln Ser Lys Thr Val
1630 1635 1640

gtc acc acg aca gtg aca gac tcc ctg acc acc acg gga ggc aca ctg 5235
Val Thr Thr Thr Val Thr Asp Ser Leu Thr Thr Thr Gly Gly Thr Leu
1645 1650 1655

gtt aca tct atg act gtg agc aaa gag tat tcc aca cga gac aaa gtg 5283
Val Thr Ser Met Thr Val Ser Lys Glu Tyr Ser Thr Arg Asp Lys Val
1660 1665 1670 1675

aaa ctg atg aaa ttt tca aga cca aag aag act cgt tca ggt aca gct 5331
Lys Leu Met Lys Phe Ser Arg Pro Lys Lys Thr Arg Ser Gly Thr Ala

1680	1685	1690	
ctg cca tcc tat aga aaa ttt gtt acc aag agc acc aag aag agc att	5379		
Leu Pro Ser Tyr Arg Lys Phe Val Thr Lys Ser Thr Lys Lys Ser Ile			
1695	1700	1705	
ttt gtt ttg cct aat gat gac tta aaa aag ttg gcc cga aaa gga gga	5427		
Phe Val Leu Pro Asn Asp Asp Leu Lys Lys Leu Ala Arg Lys Gly Gly			
1710	1715	1720	
atc cga gag gtc cct tat ttt aat tac aat gca aaa cct gct ttg gat	5475		
Ile Arg Glu Val Pro Tyr Phe Asn Tyr Asn Ala Lys Pro Ala Leu Asp			
1725	1730	1735	
ata tgg cca tat cct tct cct aga ccg acc ttt ggc atc act tgg agg	5523		
Ile Trp Pro Tyr Pro Ser Pro Arg Pro Thr Phe Gly Ile Thr Trp Arg			
1740	1745	1750	1755
tat aga ctt cag aca gta aag tcc tta gct gga gtg agc ctg atg tta	5571		
Tyr Arg Leu Gln Thr Val Lys Ser Leu Ala Gly Val Ser Leu Met Leu			
1760	1765	1770	
cgg tta ctg tgg gca agt ttg aga tgg gat gat atg gcg gcc aag gtt	5619		
Arg Leu Leu Trp Ala Ser Leu Arg Trp Asp Asp Met Ala Ala Lys Val			
1775	1780	1785	

cct cca gga gga ggg agt aca cgg aca gaa aca tcc gaa act gaa atc 5667

Pro Pro Gly Gly Gly Ser Thr Arg Thr Glu Thr Ser Glu Thr Glu Ile

1790

1795

1800

aca aca aca gaa ata att aag agg aga gat gtt ggt cct tat ggc att 5715

Thr Thr Thr Glu Ile Ile Lys Arg Arg Asp Val Gly Pro Tyr Gly Ile

1805

1810

1815

cga ttt gaa tat tgt atc agg aaa atc att tgt ccc att gga gtt cca 5763

Arg Phe Glu Tyr Cys Ile Arg Lys Ile Ile Cys Pro Ile Gly Val Pro

1820

1825

1830

1835

gaa aca cca aaa gaa acg cct aca cct cag agg aaa ggc ctt cga tca 5811

Glu Thr Pro Lys Glu Thr Pro Thr Pro Gln Arg Lys Gly Leu Arg Ser

1840

1845

1850

agt gca ctg cgg cca aag aga cca gaa acg ccc aag caa act ggc cct 5859

Ser Ala Leu Arg Pro Lys Arg Pro Glu Thr Pro Lys Gln Thr Gly Pro

1855

1860

1865

gtt att att gaa acc tgg gta gca gaa gaa gaa ctg gaa ttg tgg gag 5907

Val Ile Ile Glu Thr Trp Val Ala Glu Glu Glu Leu Glu Leu Trp Glu

1870

1875

1880

atc agg gca ttt gct gag aga gtg gag aaa gaa aag gca caa gca gtt 5955

Ile Arg Ala Phe Ala Glu Arg Val Glu Lys Glu Lys Ala Gln Ala Val

1885 1890 1895

gag caa cag gct aag aaa cga ctg gag cag cag aag ccg aca gtg att 6003
Glu Gln Gln Ala Lys Lys Arg Leu Glu Gln Gln Lys Pro Thr Val Ile

1900 1905 1910 1915

gca act tcc act act tcc cca aca agc agt aca acc agc acc atc tct 6051
Ala Thr Ser Thr Thr Ser Pro Thr Ser Ser Thr Thr Ser Thr Ile Ser

 1920 1925 1930

cca gca cag aaa gtt atg gtg gcc ccc ata agt ggc tca gtt aca act 6099
Pro Ala Gln Lys Val Met Val Ala Pro Ile Ser Gly Ser Val Thr Thr

 1935 1940 1945

gga acc aaa atg gta cta act act aaa gtt gga tct cca gct aca gta 6147
Gly Thr Lys Met Val Leu Thr Thr Lys Val Gly Ser Pro Ala Thr Val

 1950 1955 1960

aca ttc caa caa aac aag aac ttt cat caa acc ttt gct aca tgg gtt 6195
Thr Phe Gln Gln Asn Lys Asn Phe His Gln Thr Phe Ala Thr Trp Val

 1965 1970 1975

aag caa ggc cag tca aat tca ggc gtt gtt caa gta cag cag aaa gtc 6243
Lys Gln Gly Gln Ser Asn Ser Gly Val Val Gln Val Gln Gln Lys Val

1980 1985 1990 1995

ctg ggt atc att cca tca agt aca ggt acc agt cag caa acc ttt act 6291
Leu Gly Ile Ile Pro Ser Ser Thr Gly Thr Ser Gln Gln Thr Phe Thr

2000

2005

2010

tca ttc cag ccc agg aca gca aca gtc aca att agg ccc aat acc tca 6339
Ser Phe Gln Pro Arg Thr Ala Thr Val Thr Ile Arg Pro Asn Thr Ser

2015

2020

2025

ggc tct gga gga acc aca agc aat tca caa gta atc aca ggg cct cag 6387
Gly Ser Gly Gly Thr Thr Ser Asn Ser Gln Val Ile

2030

2035

2040

att cgc cct ggt atg acc gtg att aga aca cca ctc caa cag tca aca 6435
Ile Arg Pro Gly Met Thr Val Ile Arg Thr Pro Leu Gln Gln Ser Thr

2045

2050

2055

cta gga aag gca att att cga aca cct gtg atg gta cag cca ggt gct 6483
Leu Gly Lys Ala Ile Ile Arg Thr Pro Val Met Val Gln Pro Gly Ala

2060

2065

2070

2075

cct cag caa gtg atg act caa atc atc agg ggg cag cct gtc tcc act 6531
Pro Gln Gln Val Met Thr Gln Ile Ile Arg Gly Gln Pro Val Ser Thr

2080

2085

2090

gca gtc tcc gcc cct aac acg gtt tcc tca aca cct ggg cag aaa agc 6579
Ala Val Ser Ala Pro Asn Thr Val Ser Ser Thr Pro Gly Gln Lys Ser

2095	2100	2105	
tta act tca gca acg tcc act tca aat ata cag tct tca gcc tca caa 6627			
Leu Thr Ser Ala Thr Ser Thr Ser Asn Ile Gln Ser Ser Ala Ser Gln			
2110	2115	2120	
ccc cct cgc ccc caa caa gga caa gtg aag ctc acc atg gct caa ctt 6675			
Pro Pro Arg Pro Gln Gln Gly Gln Val Lys Leu Thr Met Ala Gln Leu			
2125	2130	2135	
act cag tta aca cag ggc cac ggt ggc aat caa ggt ttg aca gta gta 6723			
Thr Gln Leu Thr Gln Gly His Gly Gly Asn Gln Gly Leu Thr Val Val			
2140	2145	2150	2155
att caa gga caa ggt caa act act gga cag ttg cag ttg ata cct caa 6771			
Ile Gln Gly Gln Gly Gln Thr Thr Gly Gln Leu Gln Leu Ile Pro Gln			
2160	2165	2170	
ggg gtg act gta ctc cca ggc cca ggc cag cag cta atg caa gct gca 6819			
Gly Val Thr Val Leu Pro Gly Pro Gly Gln Gln Leu Met Gln Ala Ala			
2175	2180	2185	
atg cca aat ggt act gtt cag cga ttc ctc ttt acc cca ttg gca aca 6867			
Met Pro Asn Gly Thr Val Gln Arg Phe Leu Phe Thr Pro Leu Ala Thr			
2190	2195	2200	

aca gcc acc aca gcc agc acc acc acc acc act gtt tcc acg aca gca 6915

Thr Ala Thr Thr Ala Ser Thr Thr Thr Thr Thr Val Ser Thr Thr Ala

2205

2210

2215

gca ggt aca ggt gaa caa agg cag agt aaa ctg tca ccc cag atg cag 6963

Ala Gly Thr Gly Glu Gln Arg Gln Ser Lys Leu Ser Pro Gln Met Gln

2220

2225

2230

2235

gta cat caa gac aaa acc ctg cca cca gct cag tca tca agt gtg ggt 7011

Val His Gln Asp Lys Thr Leu Pro Pro Ala Gln Ser Ser Ser Val Gly

2240

2245

2250

cca gca aaa gcc cag cca cag act gct cag cct tca gct cgg ccc cag 7059

Pro Ala Lys Ala Gln Pro Gln Thr Ala Gln Pro Ser Ala Arg Pro Gln

2255

2260

2265

ccc caa acc cag ccc cag tcc cca gct cag cct gaa gtt cag act cag 7107

Pro Gln Thr Gln Pro Gln Ser Pro Ala Gln Pro Glu Val Gln Thr Gln

2270

2275

2280

cct gaa gtt cag acc caa aca act gtt tca tcc cat gtc cct tct gaa 7155

Pro Glu Val Gln Thr Gln Thr Thr Val Ser Ser His Val Pro Ser Glu

2285

2290

2295

gca caa ccc acc cac gca cag tca tcc aag ccc caa gtt gca gca cag 7203

Ala Gln Pro Thr His Ala Gln Ser Ser Lys Pro Gln Val Ala Ala Gln

2300	2305	2310	2315
tct cag cct caa agt aat gtc caa gga cag tct cct gtt cgt gtc caa 7251			
Ser Gln Pro Gln Ser Asn Val Gln Gly Gln Ser Pro Val Arg Val Gln			
	2320	2325	2330
agt cca tca cag act cga ata cgt cca tca act cca tcc caa ctg tct 7299			
Ser Pro Ser Gln Thr Arg Ile Arg Pro Ser Thr Pro Ser Gln Leu Ser			
	2335	2340	2345
cct gga caa caa tcc cag gtt cag act aca acc tca caa ccg att cca 7347			
Pro Gly Gln Gln Ser Gln Val Gln Thr Thr Thr Ser Gln Pro Ile Pro			
	2350	2355	2360
att caa cca cat aca tct ctt cag ata cct tcc caa ggc cag cca cag 7395			
Ile Gln Pro His Thr Ser Leu Gln Ile Pro Ser Gln Gly Gln Pro Gln			
	2365	2370	2375
tca caa ccc cag gta cag tct tca act caa act ctt tca tca gga caa 7443			
Ser Gln Pro Gln Val Gln Ser Ser Thr Gln Thr Leu Ser Ser Gly Gln			
2380	2385	2390	2395
act tta aat caa gtt agt gtt tca tcc cca tcc cgt cct cag cta caa 7491			
Thr Leu Asn Gln Val Ser Val Ser Ser Pro Ser Arg Pro Gln Leu Gln			
	2400	2405	2410

ata cag cag cca cag ccc caa gtc att gct gtg cct cag ctg caa caa 7539

Ile Gln Gln Pro Gln Pro Gln Val Ile Ala Val Pro Gln Leu Gln Gln

2415

2420

2425

caa gtc cag gtt ctc tct cag atc cag tca cag gtt gtg gct cag ata 7587

Gln Val Gln Val Leu Ser Gln Ile Gln Ser Gln Val Val Ala Gln Ile

2430

2435

2440

cag gct cag caa agt ggt gtg ccc cag caa atc aaa ctc cag tta cct 7635

Gln Ala Gln Gln Ser Gly Val Pro Gln Gln Ile Lys Leu Gln Leu Pro

2445

2450

2455

atc caa att cag caa agc agt gct gtg cag act cac cag att cag aat 7683

Ile Gln Ile Gln Gln Ser Ser Ala Val Gln Thr His Gln Ile Gln Asn

2460

2465

2470

2475

gtg gtt aca gtg cag gca gcc agt gtg caa gag cag ttg caa agg gtt 7731

Val Val Thr Val Gln Ala Ala Ser Val Gln Glu Gln Leu Gln Arg Val

2480

2485

2490

cag caa ctc agg gat cag cag caa aag aag aaa cag caa cag ata gaa 7779

Gln Gln Leu Arg Asp Gln Gln Gln Lys Lys Lys Gln Gln Gln Ile Glu

2495

2500

2505

att aac gtg aac aca ccc tcc aag ctt cta atc aaa gtt gaa atc att 7827

Ile Asn Val Asn Thr Pro Ser Lys Leu Leu Ile Lys Val Glu Ile Ile

2510	2515	2520	
cag aaa cag gtg gtg atg aag cat aat gct gta ata gaa cat tta aaa 7875			
Gln Lys Gln Val Val Met Lys His Asn Ala Val Ile Glu His Leu Lys			
2525	2530	2535	
cag aaa aag agc atg act cca gct gaa aga gaa gag aat caa aga atg 7923			
Gln Lys Lys Ser Met Thr Pro Ala Glu Arg Glu Glu Asn Gln Arg Met			
2540	2545	2550	2555
att gtc tgt aac cag gtg atg aag tat att ttg gat aag ata gat aaa 7971			
Ile Val Cys Asn Gln Val Met Lys Tyr Ile Leu Asp Lys Ile Asp Lys			
2560	2565	2570	
gaa gaa aaa cag gca gca aaa aaa cgg aag cgt gaa gag agt gtg gag 8019			
Glu Glu Lys Gln Ala Ala Lys Lys Arg Lys Arg Glu Glu Ser Val Glu			
2575	2580	2585	
cag aaa cgt agc aag cag aat gcc act aag ctg tca gct ctg ctc ttc 8067			
Gln Lys Arg Ser Lys Gln Asn Ala Thr Lys Leu Ser Ala Leu Leu Phe			
2590	2595	2600	
aag cac aaa gag cag ctc aga gcc gag atc ctg aag aag aga gca ctc 8115			
Lys His Lys Glu Gln Leu Arg Ala Glu Ile Leu Lys Lys Arg Ala Leu			
2605	2610	2615	

ctg gac aag gat ctg caa att gaa gtg cag gaa gag ctg aag aga gac 8163

Leu Asp Lys Asp Leu Gln Ile Glu Val Gln Glu Glu Leu Lys Arg Asp

2620 2625 2630 2635

ctg aaa att aag aaa gaa aaa gac ctg atg cag ttg gct cag gcc aca 8211

Leu Lys Ile Lys Lys Glu Lys Asp Leu Met Gln Leu Ala Gln Ala Thr

2640 2645 2650

gca gta gct gca ccc tgc ccc cca gtg aca cca gtt ctt cca gcc cct 8259

Ala Val Ala Ala Pro Cys Pro Pro Val Thr Pro Val Leu Pro Ala Pro

2655 2660 2665

cca gcc cct cca cct tca cct ccc cct cca cct ggt gtg caa cac aca 8307

Pro Ala Pro Pro Pro Ser Pro Pro Pro Pro Pro Gly Val Gln His Thr

2670 2675 2680

ggc ctt ctg tcc acg ccc acc tta cct gtt gct tcc cag aag agg aag 8355

Gly Leu Leu Ser Thr Pro Thr Leu Pro Val Ala Ser Gln Lys Arg Lys

2685 2690 2695

cgg gaa gag gaa aaa gac tcc agc tca aag tcc aag aaa aag aaa atg 8403

Arg Glu Glu Glu Lys Asp Ser Ser Ser Lys Ser Lys Lys Lys Lys Met

2700 2705 2710 2715

atc tct act acc tca aag gaa act aag aag gac aca aag ctt tac tgt 8451

Ile Ser Thr Thr Ser Lys Glu Thr Lys Lys Asp Thr Lys Leu Tyr Cys

2720	2725	2730	
atc tgt aaa acg cct tat gat gaa tct aaa ttt tat att ggc tgt gat 8499			
Ile Cys Lys Thr Pro Tyr Asp Glu Ser Lys Phe Tyr Ile Gly Cys Asp			
2735	2740	2745	
cgg tgt cag aat tgg tac cat ggg cgc tgc gtt ggc atc ttg caa agt 8547			
Arg Cys Gln Asn Trp Tyr His Gly Arg Cys Val Gly Ile Leu Gln Ser			
2750	2755	2760	
gag gca gag ctc att gat gag tat gtc tgt cca cag tgc cag tca aca 8595			
Glu Ala Glu Leu Ile Asp Glu Tyr Val Cys Pro Gln Cys Gln Ser Thr			
2765	2770	2775	
gag gat gcc atg aca gtg ctc acg cca cta aca gag aag gat tat gag 8643			
Glu Asp Ala Met Thr Val Leu Thr Pro Leu Thr Glu Lys Asp Tyr Glu			
2780	2785	2790	2795
ggg ttg aag agg gtg ctc cgt tcc tta cag gcc cat aag atg gcc tgg 8691			
Gly Leu Lys Arg Val Leu Arg Ser Leu Gln Ala His Lys Met Ala Trp			
2800	2805	2810	
cct ttc ctt gaa cca gta gac cct aat gat gca cca gat tat tat ggt 8739			
Pro Phe Leu Glu Pro Val Asp Pro Asn Asp Ala Pro Asp Tyr Tyr Gly			
2815	2820	2825	

gtt att aag gaa cct atg gac ctt gcc acc atg gaa gaa aga gta caa 8787

Val Ile Lys Glu Pro Met Asp Leu Ala Thr Met Glu Glu Arg Val Gln

2830

2835

2840

aga cga tat tat gaa aag ctg acg gaa ttt gtg gca gat atg acc aaa 8835

Arg Arg Tyr Tyr Glu Lys Leu Thr Glu Phe Val Ala Asp Met Thr Lys

2845

2850

2855

att ttt gat aac tgt cgt tac tac aat cca agt gac tcc cca ttt tac 8883

Ile Phe Asp Asn Cys Arg Tyr Tyr Asn Pro Ser Asp Ser Pro Phe Tyr

2860

2865

2870

2875

cag tgt gca gaa gtt ctc gaa tca ttc ttt gta cag aaa ttg aaa ggc 8931

Gln Cys Ala Glu Val Leu Glu Ser Phe Phe Val Gln Lys Leu Lys Gly

2880

2885

2890

ttc aaa gct agc agg tct cat aac aac aaa ctg cag tct aca gct tct 8979

Phe Lys Ala Ser Arg Ser His Asn Asn Lys Leu Gln Ser Thr Ala Ser

2895

2900

2905

taaagttcag cgtgttaacc taacataaaa cacagcaaga atctggttgt ctgaactatt 9039

ttaaattaag gagccagatg ttttagtca ggctatcctg acaagacttg acctaaactt 9099

cgtttttatt ggtcataaca gtccaattat attcttgcc aattttgtcc aacggacaag 9159

aaaaaagcaa agtcaacgac accattatct tgtcaagatc agatggtttt actattgtgg 9219

cagaagcgag aaaactttgt ttattgaaaa aaaaagaaaa agaaagcaag aaaaaaagat 9279

actatgggggt caagtgtaac tccatggaaa tgccacgtct gctcttcagt gaagaagctg 9339

gtttagagtc tcacagaaaa cttttgactg tatttattta ttgttgcaaa aaagacgctt 9399

ttttattgct gccctcattt gtcagctaag tattttttct tataaaatcc agccccggtt 9459

acataatac atctgtatct tatcatgatt cctgtaggta aaagtacaag acgacctcta 9519

gatgtctttt ctttctatga aaggagctgc tatgtacaca tgtgcacaca cacacaactg 9579

ggaatcaaca atgagtttat tgttcatggt agattaaaat taagcttgca taaaggttgg 9639

gctaagtgggt ccttgggcta cagactctgt tgccttgaat ataacagtac aatttgtcaa 9699

ttactctgca ccaggctaaa gtgagtaaaa tctatttgaa ggtatcttgt ttgtaaacad 9759

ttgtcagatt ctaatttttt tcttttgtat taaaattcaa ctatggatgt atatgaaaca 9819

aaataaatgg agataatttt tctcccacaa aaaaaaaaaa aaaaaa 9865

<210> 3

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Artificially
Synthesized Sequence

<400> 3

ggattatgag gggttgaaga ggg

23

<210> 4

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Artificially
Synthesized Sequence

<400> 4

aaggcaacag agtctgtagc ccaa

24

<210> 5

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Artificially
Synthesized Sequence

<400> 5

cctcagctgc aacaagtcc

19

<210> 6

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Artificially
Synthesized Sequence

<400> 6

gcactgcttt gctgaatttg ga

22

<210> 7

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Artificially
Synthesized Sequence

<400> 7

aagatgttgt cttggagccg t

21

<210> 8

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Artificially
Synthesized Sequence

<400> 8

ttttttacca ttgcttcag tccc

24

<210> 9

<211> 9700

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (472)..(8814)

<400> 9

agccgccact ggtccggcc ctcccgctca gtttccctt ctcccgccgc ctgggtcca 60

acaagagggg ccggcggggc aggccgacca agcagcccgc ggctccgct gcggagcgt 120

ggccccggc cccgccgccc cgcgcgccc cgtccggacc catcgggggc tccctcgcc 180

gatacgcggt agtagccggg gcaggtgggc agccgccagg ctgaggtggc gccaagacg 240

cggtgagct cgcgcagggt gggcagcagt agccggagga agccgccgc gccgccgccc 300

gccccccca gcaccagcgc cccgggcccgg ggggggcgag gaggcggggg cggcacgacg 360

ggggcgggg gcggcgggc ccacctgtcc cggaccaccg cggcccgag ggccgtcaac 420

aaagtggtgt acgatgacca cgagagcgag gcggtggagg aagaggagga c atg gtc 477

Met Val

1

tcc gag gag gag gag gag gac ggc gac gcc gag gag acc cag gat 525

Ser Glu Glu Glu Glu Glu Glu Asp Gly Asp Ala Glu Glu Thr Gln Asp

5

10

15

tct gag gac gac gag gag gat gag atg gaa gag gac gac gat gac tcc 573

Ser Glu Asp Asp Glu Glu Asp Glu Met Glu Glu Asp Asp Asp Asp Ser

20

25

30

gat tat ccg gag gag atg gaa gac gac gac gac gac gcc agt tac tgc 621

Asp Tyr Pro Glu Glu Met Glu Asp Asp Asp Asp Ala Ser Tyr Cys

35

40

45

50

acg gaa agc agc ttc agg agc cat agt acc tac agc agc act cca ggt 669

Thr Glu Ser Ser Phe Arg Ser His Ser Thr Tyr Ser Ser Thr Pro Gly

55

60

65

agg cga aaa cca aga gta cat cgg cct cgt tct cct ata ttg gaa gaa 717

Arg Arg Lys Pro Arg Val His Arg Pro Arg Ser Pro Ile Leu Glu Glu

70

75

80

aaa gac atc ccg ccc ctt gaa ttt ccc aag tcc tct gag gat tta atg 765

Lys Asp Ile Pro Pro Leu Glu Phe Pro Lys Ser Ser Glu Asp Leu Met

85

90

95

gtg cct aat gag cat ata atg aat gtc att gcc att tac gag gta ctg 813

Val Pro Asn Glu His Ile Met Asn Val Ile Ala Ile Tyr Glu Val Leu

100

105

110

cgg aac ttt ggc act gtt ttg aga tta tct cct ttt cgc ttt gag gac 861
Arg Asn Phe Gly Thr Val Leu Arg Leu Ser Pro Phe Arg Phe Glu Asp
115 120 125 130

ttt tgt gca gct ctg gtg agc caa gag cag tgc aca ctc atg gca gag 909
Phe Cys Ala Ala Leu Val Ser Gln Glu Gln Cys Thr Leu Met Ala Glu
135 140 145

atg cat gtt gtg ctt ttg aaa gca gtt ctg cgt gaa gaa gac act tcc 957
Met His Val Val Leu Leu Lys Ala Val Leu Arg Glu Glu Asp Thr Ser
150 155 160

aat act acc ttt gga cct gct gat ctg aaa gat agc gtt aat tcc aca 1005
Asn Thr Thr Phe Gly Pro Ala Asp Leu Lys Asp Ser Val Asn Ser Thr
165 170 175

ctg tat ttc ata gat ggg atg acg tgg cca gag gtg ctg cgg gtg tac 1053
Leu Tyr Phe Ile Asp Gly Met Thr Trp Pro Glu Val Leu Arg Val Tyr
180 185 190

tgt gag agt gat aag gag tac cat cac gtt ctt cct tac caa gag gca 1101
Cys Glu Ser Asp Lys Glu Tyr His His Val Leu Pro Tyr Gln Glu Ala
195 200 205 210

gag gac tac cca tat gga cca gta gag aac aag atc aaa gtt cta cag 1149

Glu Asp Tyr Pro Tyr Gly Pro Val Glu Asn Lys Ile Lys Val Leu Gln
215 220 225

ttt cta gtc gat cag ttt ctt aca aca aat att gct cga gag gaa ttg 1197
Phe Leu Val Asp Gln Phe Leu Thr Thr Asn Ile Ala Arg Glu Glu Leu
230 235 240

atg tct gaa ggg gtg ata cag tat gat gac cat tgt agg gtt tgt cac 1245
Met Ser Glu Gly Val Ile Gln Tyr Asp Asp His Cys Arg Val Cys His
245 250 255

aaa ctt ggg gat ttg ctt tgc tgt gag aca tgt tca gca gta tac cat 1293
Lys Leu Gly Asp Leu Leu Cys Cys Glu Thr Cys Ser Ala Val Tyr His
260 265 270

ttg gaa tgt gtg aag cca cct ctt gag gag gtg cca gag gac gag tgg 1341
Leu Glu Cys Val Lys Pro Pro Leu Glu Glu Val Pro Glu Asp Glu Trp
275 280 285 290

cag tgt gaa gtc tgt gta gca cac aag gtg cct ggt gtg act gac tgt 1389
Gln Cys Glu Val Cys Val Ala His Lys Val Pro Gly Val Thr Asp Cys
295 300 305

gtt gct gaa atc caa aaa aat aaa cca tat att cga cat gaa cct att 1437
Val Ala Glu Ile Gln Lys Asn Lys Pro Tyr Ile Arg His Glu Pro Ile
310 315 320

gga tat gat aga agt cgg agg aaa tac tgg ttc ttg aac cga aga ctc 1485

Gly Tyr Asp Arg Ser Arg Arg Lys Tyr Trp Phe Leu Asn Arg Arg Leu

325

330

335

ata ata gaa gaa gat aca gaa aat gaa aat gaa aag aaa att tgg tat 1533

Ile Ile Glu Glu Asp Thr Glu Asn Glu Asn Glu Lys Lys Ile Trp Tyr

340

345

350

tac agc aca aag gtc caa ctt gca gaa tta att gac tgt cta gac aaa 1581

Tyr Ser Thr Lys Val Gln Leu Ala Glu Leu Ile Asp Cys Leu Asp Lys

355

360

365

370

gat tat tgg gaa gca gaa ctc tgc aaa att cta gaa gaa atg cgt gaa 1629

Asp Tyr Trp Glu Ala Glu Leu Cys Lys Ile Leu Glu Glu Met Arg Glu

375

380

385

gaa atc cac cga cac atg gac ata act gaa gac ctg acc aat aag gct 1677

Glu Ile His Arg His Met Asp Ile Thr Glu Asp Leu Thr Asn Lys Ala

390

395

400

cgg ggc agt aac aaa tcc ttt ctg gcg gca gct aat gaa gaa att ttg 1725

Arg Gly Ser Asn Lys Ser Phe Leu Ala Ala Ala Asn Glu Glu Ile Leu

405

410

415

gag tcc ata aga gcc aaa aag gga gac att gat aat gtt aaa agc cca 1773

Glu Ser Ile Arg Ala Lys Lys Gly Asp Ile Asp Asn Val Lys Ser Pro

420

425

430

gaa gaa aca gaa aaa gac aag aat gag act gag aat gac tct aaa gat 1821

Glu Glu Thr Glu Lys Asp Lys Asn Glu Thr Glu Asn Asp Ser Lys Asp

435

440

445

450

gct gag aaa aac aga gaa gaa ttt gaa gac cag tcc ctt gaa aaa gac 1869

Ala Glu Lys Asn Arg Glu Glu Phe Glu Asp Gln Ser Leu Glu Lys Asp

455

460

465

agt gac gac aaa aca cca gat gat gac cct gag caa gga aaa tct gag 1917

Ser Asp Asp Lys Thr Pro Asp Asp Asp Pro Glu Gln Gly Lys Ser Glu

470

475

480

gta ggt gat ttc aaa tcg gag aag tcc aac ggg gag cta agt gaa tct 1965

Val Gly Asp Phe Lys Ser Glu Lys Ser Asn Gly Glu Leu Ser Glu Ser

485

490

495

cct gga gct gga aaa gga gca tct ggc tca act cga atc atc acc aga 2013

Pro Gly Ala Gly Lys Gly Ala Ser Gly Ser Thr Arg Ile Ile Thr Arg

500

505

510

ttg cgg aat cca gat agc aaa ctt agt cag ctg aag agc cag cag gtg 2061

Leu Arg Asn Pro Asp Ser Lys Leu Ser Gln Leu Lys Ser Gln Gln Val

515

520

525

530

gca gcc gct gca cat gaa gca aat aaa tta ttt aag gag ggc aaa gag 2109

Ala Ala Ala Ala His Glu Ala Asn Lys Leu Phe Lys Glu Gly Lys Glu

535

540

545

gta ctg gta gtt aac tct caa gga gaa att tca cgg ttg agc acc aaa 2157

Val Leu Val Val Asn Ser Gln Gly Glu Ile Ser Arg Leu Ser Thr Lys

550

555

560

aag gaa gtg atc atg aaa gga aat atc aac aat tat ttt aaa ttg ggt 2205

Lys Glu Val Ile Met Lys Gly Asn Ile Asn Asn Tyr Phe Lys Leu Gly

565

570

575

caa gaa ggg aag tat cgc gtc tac cac aat caa tac tcc acc aat tca 2253

Gln Glu Gly Lys Tyr Arg Val Tyr His Asn Gln Tyr Ser Thr Asn Ser

580

585

590

ttt gct ttg aat aag cac cag cac aga gaa gac cat gat aag aga agg 2301

Phe Ala Leu Asn Lys His Gln His Arg Glu Asp His Asp Lys Arg Arg

595

600

605

610

cat ctt gca cat aag ttc tgt ctg act cca gca gga gag ttc aaa tgg 2349

His Leu Ala His Lys Phe Cys Leu Thr Pro Ala Gly Glu Phe Lys Trp

615

620

625

aac ggt tct gtc cat ggg tcc aaa gtt ctt acc ata tct act ctg aga 2397

Asn Gly Ser Val His Gly Ser Lys Val Leu Thr Ile Ser Thr Leu Arg

630

635

640

ctg act atc acc caa tta gaa aac aac atc cct tca tcc ttt ctt cat 2445

Leu Thr Ile Thr Gln Leu Glu Asn Asn Ile Pro Ser Ser Phe Leu His

645

650

655

ccc aac tgg gca tca cat agg gca aat tgg atc aag gca gtt cag atg 2493

Pro Asn Trp Ala Ser His Arg Ala Asn Trp Ile Lys Ala Val Gln Met

660

665

670

tgt agc aaa ccc aga gaa ttt gca ttg gct tta gcc att ttg gag tgt 2541

Cys Ser Lys Pro Arg Glu Phe Ala Leu Ala Leu Ala Ile Leu Glu Cys

675

680

685

690

gca gtt aaa cca gtt gtg atg cta cca ata tgg cga gaa ttt tta gga 2589

Ala Val Lys Pro Val Val Met Leu Pro Ile Trp Arg Glu Phe Leu Gly

695

700

705

cat acc agg tta cac cgg atg aca tca att gaa aga gaa gaa aag gag 2637

His Thr Arg Leu His Arg Met Thr Ser Ile Glu Arg Glu Glu Lys Glu

710

715

720

aaa gtc aaa aaa aaa gag aag aaa cag gaa gaa gaa gaa acg atg cag 2685

Lys Val Lys Lys Lys Glu Lys Lys Gln Glu Glu Glu Glu Thr Met Gln

725

730

735

caa gcg aca tgg gta aaa tac aca ttt cca gtt aag cat cag gtt tgg 2733

Gln Ala Thr Trp Val Lys Tyr Thr Phe Pro Val Lys His Gln Val Trp

740

745

750

aaa caa aaa ggt gaa gag tac aga gtg aca gga tat ggt ggt tgg agc 2781

Lys Gln Lys Gly Glu Glu Tyr Arg Val Thr Gly Tyr Gly Gly Trp Ser

755

760

765

770

tgg att agt aaa act cat gtt tat agg ttt gtt cct aaa ttg cca ggc 2829

Trp Ile Ser Lys Thr His Val Tyr Arg Phe Val Pro Lys Leu Pro Gly

775

780

785

aat act aat gtg aat tac aga aag tcg tta gaa gga acc aaa aat aat 2877

Asn Thr Asn Val Asn Tyr Arg Lys Ser Leu Glu Gly Thr Lys Asn Asn

790

795

800

atg gat gaa aat atg gat gag tca gat aaa aga aaa tgt tca cga agt 2925

Met Asp Glu Asn Met Asp Glu Ser Asp Lys Arg Lys Cys Ser Arg Ser

805

810

815

cca aaa aaa ata aaa ata gag cct gat tct gaa aaa gat gag gta aaa 2973

Pro Lys Lys Ile Lys Ile Glu Pro Asp Ser Glu Lys Asp Glu Val Lys

820

825

830

ggt tca gat gct gca aaa gga gca gac caa aat gaa atg gat atc tca 3021

Gly Ser Asp Ala Ala Lys Gly Ala Asp Gln Asn Glu Met Asp Ile Ser

835 840 845 850

aag att act gag aag aag gac caa gat gtg aag gag ctc tta gat tct 3069

Lys Ile Thr Glu Lys Lys Asp Gln Asp Val Lys Glu Leu Leu Asp Ser

855 860 865

gac agt gat aaa ccc tgc aag gaa gaa cca atg gaa gta gac gat gac 3117

Asp Ser Asp Lys Pro Cys Lys Glu Glu Pro Met Glu Val Asp Asp Asp

870 875 880

atg aaa aca gag tca cat gta aat tgt cag gag agt tct caa gta gat 3165

Met Lys Thr Glu Ser His Val Asn Cys Gln Glu Ser Ser Gln Val Asp

885 890 895

gtg gtc aat gtt agt gag ggt ttt cat cta agg act agt tac aaa aag 3213

Val Val Asn Val Ser Glu Gly Phe His Leu Arg Thr Ser Tyr Lys Lys

900 905 910

aaa aca aaa tca tcc aaa cta gat gga ctt ctt gaa agg aga att aaa 3261

Lys Thr Lys Ser Ser Lys Leu Asp Gly Leu Leu Glu Arg Arg Ile Lys

915 920 925 930

cag ttt aca ctg gaa gaa aaa cag cga ctc gaa aaa atc aag ttg gag 3309

Gln Phe Thr Leu Glu Glu Lys Gln Arg Leu Glu Lys Ile Lys Leu Glu

935 940 945

ggt gga att aag ggt ata gga aag act tct aca aat tct tca aaa aat 3357

Gly Gly Ile Lys Gly Ile Gly Lys Thr Ser Thr Asn Ser Ser Lys Asn

950

955

960

ctc tct gaa tca cca gta ata acg aaa gca aaa gaa ggg tgt cag agt 3405

Leu Ser Glu Ser Pro Val Ile Thr Lys Ala Lys Glu Gly Cys Gln Ser

965

970

975

gac tcg atg aga caa gaa cag agc cca aat gca aat aat gat caa cct 3453

Asp Ser Met Arg Gln Glu Gln Ser Pro Asn Ala Asn Asn Asp Gln Pro

980

985

990

gag gac ttg att cag gga tgt tca caa agt gat tcc tca gtt ctt aga 3501

Glu Asp Leu Ile Gln Gly Cys Ser Gln Ser Asp Ser Ser Val Leu Arg

995

1000

1005

1010

atg agt gat cct agt cat acc aca aac aaa ctt tat cca aaa gat cga 3549

Met Ser Asp Pro Ser His Thr Thr Asn Lys Leu Tyr Pro Lys Asp Arg

1015

1020

1025

gtg tta gat gat gtc tcc att cgg agc cca gaa aca aaa tgt ccg aaa 3597

Val Leu Asp Asp Val Ser Ile Arg Ser Pro Glu Thr Lys Cys Pro Lys

1030

1035

1040

caa aat tcc att gaa aat gac ata gaa gaa aaa gtc tct gac ctt gcc 3645

Gln Asn Ser Ile Glu Asn Asp Ile Glu Glu Lys Val Ser Asp Leu Ala

1045

1050

1055

agt aga ggc cag gaa ccc act aag agt aaa acc aaa gga aat gat ttt 3693

Ser Arg Gly Gln Glu Pro Thr Lys Ser Lys Thr Lys Gly Asn Asp Phe

1060

1065

1070

ttc atc gat gac tct aaa cta gcc agt gca gat gat att ggt act ttg 3741

Phe Ile Asp Asp Ser Lys Leu Ala Ser Ala Asp Asp Ile Gly Thr Leu

1075

1080

1085

1090

atc tgt aag aac aaa aaa ccg ctc ata cag gag gaa agt gac acc att 3789

Ile Cys Lys Asn Lys Lys Pro Leu Ile Gln Glu Glu Ser Asp Thr Ile

1095

1100

1105

gtt tct tct tcc aag agt gct tta cat tca tca gtg cct aaa agt acc 3837

Val Ser Ser Ser Lys Ser Ala Leu His Ser Ser Val Pro Lys Ser Thr

1110

1115

1120

aat gac aga gat gcc aca cct ctg tca aga gca atg gac ttt gaa gga 3885

Asn Asp Arg Asp Ala Thr Pro Leu Ser Arg Ala Met Asp Phe Glu Gly

1125

1130

1135

aaa ctg gga tgt gac tct gaa tct aat agc act ttg gaa aat agt tct 3933

Lys Leu Gly Cys Asp Ser Glu Ser Asn Ser Thr Leu Glu Asn Ser Ser

1140

1145

1150

gat acc gtg tct att cag gat agc agt gaa gaa gat atg att gtt cag 3981
Asp Thr Val Ser Ile Gln Asp Ser Ser Glu Glu Asp Met Ile Val Gln
1155 1160 1165 1170

aat agc aat gaa agc att tct gaa cag ttc aga act cga gaa caa gat 4029
Asn Ser Asn Glu Ser Ile Ser Glu Gln Phe Arg Thr Arg Glu Gln Asp
1175 1180 1185

gtt gaa gtc ttg gag ccg tta aag tgt gag ttg gtt tct ggt gag tcc 4077
Val Glu Val Leu Glu Pro Leu Lys Cys Glu Leu Val Ser Gly Glu Ser
1190 1195 1200

act gga aac tgt gag gac agg ctg ccg gtc aag ggg act gaa gca aat 4125
Thr Gly Asn Cys Glu Asp Arg Leu Pro Val Lys Gly Thr Glu Ala Asn
1205 1210 1215

ggt aaa aaa cca agt cag cag aag aaa tta gag gag aga cca gtt aat 4173
Gly Lys Lys Pro Ser Gln Gln Lys Lys Leu Glu Glu Arg Pro Val Asn
1220 1225 1230

aaa tgt agt gat caa ata aag cta aaa aat acc act gac aaa aag aat 4221
Lys Cys Ser Asp Gln Ile Lys Leu Lys Asn Thr Thr Asp Lys Lys Asn
1235 1240 1245 1250

aat gaa aat cga gag tct gaa aag aaa gga cag aga aca agt aca ttt 4269

Asn Glu Asn Arg Glu Ser Glu Lys Lys Gly Gln Arg Thr Ser Thr Phe

1255

1260

1265

caa ata aat gga aaa gat aat aaa ccc aaa ata tat ttg aaa ggt gaa 4317

Gln Ile Asn Gly Lys Asp Asn Lys Pro Lys Ile Tyr Leu Lys Gly Glu

1270

1275

1280

tgc ttg aaa gaa att tct gag agt aga gta gta agt ggt aat gtt gaa 4365

Cys Leu Lys Glu Ile Ser Glu Ser Arg Val Val Ser Gly Asn Val Glu

1285

1290

1295

cca aag gtt aat aat ata aat aaa ata atc cct gag aat gat att aaa 4413

Pro Lys Val Asn Asn Ile Asn Lys Ile Ile Pro Glu Asn Asp Ile Lys

1300

1305

1310

tca ttg act gtt aaa gaa tct gct ata agg cca ttc att aat ggt gat 4461

Ser Leu Thr Val Lys Glu Ser Ala Ile Arg Pro Phe Ile Asn Gly Asp

1315

1320

1325

1330

gtc atc atg gaa gat ttt aat gaa aga aac agc tcc gaa aca aaa tcg 4509

Val Ile Met Glu Asp Phe Asn Glu Arg Asn Ser Ser Glu Thr Lys Ser

1335

1340

1345

cat ttg ctg agt tct tca gat gct gaa ggt aac tac cga gat agc ctt 4557

His Leu Leu Ser Ser Ser Asp Ala Glu Gly Asn Tyr Arg Asp Ser Leu

1350

1355

1360

gag acc ctg cca tca acc aaa gag tct gac agt aca cag acg acc aca 4605

Glu Thr Leu Pro Ser Thr Lys Glu Ser Asp Ser Thr Gln Thr Thr Thr

1365

1370

1375

ccc tca gca tct tgt cca gaa agc aat tca gtt aat cag gta gaa gat 4653

Pro Ser Ala Ser Cys Pro Glu Ser Asn Ser Val Asn Gln Val Glu Asp

1380

1385

1390

atg gaa ata gaa acc tca gaa gtt aag aaa gtt act tca tca cct att 4701

Met Glu Ile Glu Thr Ser Glu Val Lys Lys Val Thr Ser Ser Pro Ile

1395

1400

1405

1410

act tct gaa gag gaa tct aat ctc agt aat gac ttt att gat gaa aat 4749

Thr Ser Glu Glu Glu Ser Asn Leu Ser Asn Asp Phe Ile Asp Glu Asn

1415

1420

1425

ggt ctg ccc atc aac aaa aat gaa aat gtc aat gga gaa tct aaa aga 4797

Gly Leu Pro Ile Asn Lys Asn Glu Asn Val Asn Gly Glu Ser Lys Arg

1430

1435

1440

aaa acc gtc atc aca gaa gtc acc acg atg acc tcc aca gtg gcc aca 4845

Lys Thr Val Ile Thr Glu Val Thr Thr Met Thr Ser Thr Val Ala Thr

1445

1450

1455

gaa tca aaa act gtg atc aag gta gaa aaa ggc gat aag caa act gtg 4893

Glu Ser Lys Thr Val Ile Lys Val Glu Lys Gly Asp Lys Gln Thr Val
1460 1465 1470

gtt tct tcc aca gaa aat tgt gca aaa tcc act gtc aca acc acc act 4941
Val Ser Ser Thr Glu Asn Cys Ala Lys Ser Thr Val Thr Thr Thr Thr
1475 1480 1485 1490

aca aca gtg acc aag ctt tcc aca ccc tcc aca ggc ggc agt gtg gac 4989
Thr Thr Val Thr Lys Leu Ser Thr Pro Ser Thr Gly Gly Ser Val Asp
1495 1500 1505

atc atc tct gta aag gag cag agc aaa acc gtg gtc acc acg aca gtg 5037
Ile Ile Ser Val Lys Glu Gln Ser Lys Thr Val Val Thr Thr Thr Val
1510 1515 1520

aca gac tcc ctg acc acc acg gga ggc aca ctg gtt aca tct atg act 5085
Thr Asp Ser Leu Thr Thr Thr Gly Gly Thr Leu Val Thr Ser Met Thr
1525 1530 1535

gtg agc aaa gag tat tcc aca cga gac aaa gtg aaa ctg atg aaa ttt 5133
Val Ser Lys Glu Tyr Ser Thr Arg Asp Lys Val Lys Leu Met Lys Phe
1540 1545 1550

tca aga cca aag aag act cgt tca ggt aca gct ctg cca tcc tat aga 5181
Ser Arg Pro Lys Lys Thr Arg Ser Gly Thr Ala Leu Pro Ser Tyr Arg
1555 1560 1565 1570

aaa ttt gtt acc aag agc acc aag aag agc att ttt gtt ttg cct aat 5229

Lys Phe Val Thr Lys Ser Thr Lys Lys Ser Ile Phe Val Leu Pro Asn

1575

1580

1585

gat gac tta aaa aag ttg gcc cga aaa gga gga atc cga gag gtc cct 5277

Asp Asp Leu Lys Lys Leu Ala Arg Lys Gly Gly Ile Arg Glu Val Pro

1590

1595

1600

tat ttt aat tac aat gca aaa cct gct ttg gat ata tgg cca tat cct 5325

Tyr Phe Asn Tyr Asn Ala Lys Pro Ala Leu Asp Ile Trp Pro Tyr Pro

1605

1610

1615

tct cct aga ccg acc ttt ggc atc act tgg agg tat aga ctt cag aca 5373

Ser Pro Arg Pro Thr Phe Gly Ile Thr Trp Arg Tyr Arg Leu Gln Thr

1620

1625

1630

gta aag tcc tta gct gga gtg agc ctg atg tta cgg tta ctg tgg gca 5421

Val Lys Ser Leu Ala Gly Val Ser Leu Met Leu Arg Leu Leu Trp Ala

1635

1640

1645

1650

agt ttg aga tgg gat gat atg gcg gcc aag gtt cct cca gga gga ggg 5469

Ser Leu Arg Trp Asp Asp Met Ala Ala Lys Val Pro Pro Gly Gly Gly

1655

1660

1665

agt aca cgg aca gaa aca tcc gaa act gaa atc aca aca aca gaa ata 5517

Ser Thr Arg Thr Glu Thr Ser Glu Thr Glu Ile Thr Thr Thr Glu Ile

1670

1675

1680

att aag agg aga gat gtt ggt cct tat ggc att cga ttt gaa tat tgt 5565

Ile Lys Arg Arg Asp Val Gly Pro Tyr Gly Ile Arg Phe Glu Tyr Cys

1685

1690

1695

atc agg aaa atc att tgt ccc att gga gtt cca gaa aca cca aaa gaa 5613

Ile Arg Lys Ile Ile Cys Pro Ile Gly Val Pro Glu Thr Pro Lys Glu

1700

1705

1710

acg cct aca cct cag agg aaa ggc ctt cga tca agt gca ctg cgg cca 5661

Thr Pro Thr Pro Gln Arg Lys Gly Leu Arg Ser Ser Ala Leu Arg Pro

1715

1720

1725

1730

aag aga cca gaa acg ccc aag caa act ggc cct gtt att att gaa acc 5709

Lys Arg Pro Glu Thr Pro Lys Gln Thr Gly Pro Val Ile Ile Glu Thr

1735

1740

1745

tgg gta gca gaa gaa gaa ctg gaa ttg tgg gag atc agg gca ttt gct 5757

Trp Val Ala Glu Glu Glu Leu Glu Leu Trp Glu Ile Arg Ala Phe Ala

1750

1755

1760

gag aga gtg gag aaa gaa aag gca caa gca gtt gag caa cag gct aag 5805

Glu Arg Val Glu Lys Glu Lys Ala Gln Ala Val Glu Gln Gln Ala Lys

1765

1770

1775

aaa cga ctg gag cag cag aag ccg aca gtg att gca act tcc act act 5853

Lys Arg Leu Glu Gln Gln Lys Pro Thr Val Ile Ala Thr Ser Thr Thr

1780

1785

1790

tcc cca aca agc agt aca acc agc acc atc tct cca gca cag aaa gtt 5901

Ser Pro Thr Ser Ser Thr Thr Ser Thr Ile Ser Pro Ala Gln Lys Val

1795

1800

1805

1810

atg gtg gcc ccc ata agt ggc tca gtt aca act gga acc aaa atg gta 5949

Met Val Ala Pro Ile Ser Gly Ser Val Thr Thr Gly Thr Lys Met Val

1815

1820

1825

cta act act aaa gtt gga tct cca gct aca gta aca ttc caa caa aac 5997

Leu Thr Thr Lys Val Gly Ser Pro Ala Thr Val Thr Phe Gln Gln Asn

1830

1835

1840

aag aac ttt cat caa acc ttt gct aca tgg gtt aag caa ggc cag tca 6045

Lys Asn Phe His Gln Thr Phe Ala Thr Trp Val Lys Gln Gly Gln Ser

1845

1850

1855

aat tca ggc gtt gtt caa gta cag cag aaa gtc ctg ggt atc att cca 6093

Asn Ser Gly Val Val Gln Val Gln Gln Lys Val Leu Gly Ile Ile Pro

1860

1865

1870

tca agt aca ggt acc agt cag caa acc ttt act tca ttc cag ccc agg 6141

Ser Ser Thr Gly Thr Ser Gln Gln Thr Phe Thr Ser Phe Gln Pro Arg
 1875 1880 1885 1890

aca gca aca gtc aca att agg ccc aat acc tca ggc tct gga gga acc 6189
 Thr Ala Thr Val Thr Ile Arg Pro Asn Thr Ser Gly Ser Gly Gly Thr
 1895 1900 1905

aca agc aat tca caa gta atc aca ggg cct cag att cgc cct ggt atg 6237
 Thr Ser Asn Ser Gln Val Ile Thr Gly Pro Gln Ile Arg Pro Gly Met
 1910 1915 1920

acc gtg att aga aca cca ctc caa cag tca aca cta gga aag gca att 6285
 Thr Val Ile Arg Thr Pro Leu Gln Gln Ser Thr Leu Gly Lys Ala Ile
 1925 1930 1935

att cga aca cct gtg atg gta cag cca ggt gct cct cag caa gtg atg 6333
 Ile Arg Thr Pro Val Met Val Gln Pro Gly Ala Pro Gln Gln Val Met
 1940 1945 1950

act caa atc atc agg ggg cag cct gtc tcc act gca gtc tcc gcc cct 6381
 Thr Gln Ile Ile Arg Gly Gln Pro Val Ser Thr Ala Val Ser Ala Pro
 1955 1960 1965 1970

aac acg gtt tcc tca aca cct ggg cag aaa agc tta act tca gca acg 6429
 Asn Thr Val Ser Ser Thr Pro Gly Gln Lys Ser Leu Thr Ser Ala Thr
 1975 1980 1985

tcc act tca aat ata cag tct tca gcc tca caa ccc cct cgc ccc caa 6477

Ser Thr Ser Asn Ile Gln Ser Ser Ala Ser Gln Pro Pro Arg Pro Gln

1990

1995

2000

caa gga caa gtg aag ctc acc atg gct caa ctt act cag tta aca cag 6525

Gln Gly Gln Val Lys Leu Thr Met Ala Gln Leu Thr Gln Leu Thr Gln

2005

2010

2015

ggc cac ggt ggc aat caa ggt ttg aca gta gta att caa gga caa ggt 6573

Gly His Gly Gly Asn Gln Gly Leu Thr Val Val Ile Gln Gly Gln Gly

2020

2025

2030

caa act act gga cag ttg cag ttg ata cct caa ggg gtg act gta ctc 6621

Gln Thr Thr Gly Gln Leu Gln Leu Ile Pro Gln Gly Val Thr Val Leu

2035

2040

2045

2050

cca ggc cca ggc cag cag cta atg caa gct gca atg cca aat ggt act 6669

Pro Gly Pro Gly Gln Gln Leu Met Gln Ala Ala Met Pro Asn Gly Thr

2055

2060

2065

gtt cag cga ttc ctc ttt acc cca ttg gca aca aca gcc acc aca gcc 6717

Val Gln Arg Phe Leu Phe Thr Pro Leu Ala Thr Thr Ala Thr Thr Ala

2070

2075

2080

agc acc acc acc acc act gtt tcc acg aca gca gca ggt aca ggt gaa 6765

Ser Thr Thr Thr Thr Thr Val Ser Thr Thr Ala Ala Gly Thr Gly Glu

2085

2090

2095

caa agg cag agt aaa ctg tca ccc cag atg cag gta cat caa gac aaa 6813

Gln Arg Gln Ser Lys Leu Ser Pro Gln Met Gln Val His Gln Asp Lys

2100

2105

2110

acc ctg cca cca gct cag tca tca agt gtg ggt cca gca aaa gcc cag 6861

Thr Leu Pro Pro Ala Gln Ser Ser Ser Val Gly Pro Ala Lys Ala Gln

2115

2120

2125

2130

cca cag act gct cag cct tca gct cgg ccc cag ccc caa acc cag ccc 6909

Pro Gln Thr Ala Gln Pro Ser Ala Arg Pro Gln Pro Gln Thr Gln Pro

2135

2140

2145

cag tcc cca gct cag cct gaa gtt cag act cag cct gaa gtt cag acc 6957

Gln Ser Pro Ala Gln Pro Glu Val Gln Thr Gln Pro Glu Val Gln Thr

2150

- 2155

2160

caa aca act gtt tca tcc cat gtc cct tct gaa gca caa ccc acc cac 7005

Gln Thr Thr Val Ser Ser His Val Pro Ser Glu Ala Gln Pro Thr His

2165

2170

2175

gca cag tca tcc aag ccc caa gtt gca gca cag tct cag cct caa agt 7053

Ala Gln Ser Ser Lys Pro Gln Val Ala Ala Gln Ser Gln Pro Gln Ser

2180

2185

2190

aat gtc caa gga cag tct cct gtt cgt gtc caa agt cca tca cag act 7101
Asn Val Gln Gly Gln Ser Pro Val Arg Val Gln Ser Pro Ser Gln Thr
2195 2200 2205 2210

cga ata cgt cca tca act cca tcc caa ctg tct cct gga caa caa tcc 7149
Arg Ile Arg Pro Ser Thr Pro Ser Gln Leu Ser Pro Gly Gln Gln Ser
2215 2220 2225

cag gtt cag act aca acc tca caa ccg att cca att caa cca cat aca 7197
Gln Val Gln Thr Thr Thr Ser Gln Pro Ile Pro Ile Gln Pro His Thr
2230 2235 2240

tct ctt cag ata cct tcc caa ggc cag cca cag tca caa ccc cag gta 7245
Ser Leu Gln Ile Pro Ser Gln Gly Gln Pro Gln Ser Gln Pro Gln Val
2245 2250 2255

cag tct tca act caa act ctt tca tca gga caa act tta aat caa gtt 7293
Gln Ser Ser Thr Gln Thr Leu Ser Ser Gly Gln Thr Leu Asn Gln Val
2260 2265 2270

agt gtt tca tcc cca tcc cgt cct cag cta caa ata cag cag cca cag 7341
Ser Val Ser Ser Pro Ser Arg Pro Gln Leu Gln Ile Gln Gln Pro Gln
2275 2280 2285 2290

ccc caa gtc att gct gtg cct cag ctg caa caa caa gtc cag gtt ctc 7389

Pro Gln Val Ile Ala Val Pro Gln Leu Gln Gln Gln Val Gln Val Leu

2295

2300

2305

tct cag atc cag tca cag gtt gtg gct cag ata cag gct cag caa agt 7437

Ser Gln Ile Gln Ser Gln Val Val Ala Gln Ile Gln Ala Gln Gln Ser

2310

2315

2320

ggt gtg ccc cag caa atc aaa ctc cag tta cct atc caa att cag caa 7485

Gly Val Pro Gln Gln Ile Lys Leu Gln Leu Pro Ile Gln Ile Gln Gln

2325

2330

2335

agc agt gct gtg cag act cac cag att cag aat gtg gtt aca gtg cag 7533

Ser Ser Ala Val Gln Thr His Gln Ile Gln Asn Val Val Thr Val Gln

2340

2345

2350

gca gcc agt gtg caa gag cag ttg caa agg gtt cag caa ctc agg gat 7581

Ala Ala Ser Val Gln Glu Gln Leu Gln Arg Val Gln Gln Leu Arg Asp

2355

2360

2365

2370

cag cag caa aag aag aaa cag caa cag ata gaa att aag cgt gaa cac 7629

Gln Gln Gln Lys Lys Lys Gln Gln Gln Ile Glu Ile Lys Arg Glu His

2375

2380

2385

acc ctc caa gct tct aat caa agt gaa atc att cag aaa cag gtg gtg 7677

Thr Leu Gln Ala Ser Asn Gln Ser Glu Ile Ile Gln Lys Gln Val Val

2390

2395

2400

atg aag cat aat gct gta ata gaa cat tta aaa cag aaa aag agc atg 7725

Met Lys His Asn Ala Val Ile Glu His Leu Lys Gln Lys Lys Ser Met

2405

2410

2415

act cca gct gaa aga gaa gag aat caa aga atg att gtc tgt aac cag 7773

Thr Pro Ala Glu Arg Glu Glu Asn Gln Arg Met Ile Val Cys Asn Gln

2420

2425

2430

gtg atg aag tat att ttg gat aag ata gat aaa gaa gaa aaa cag gca 7821

Val Met Lys Tyr Ile Leu Asp Lys Ile Asp Lys Glu Glu Lys Gln Ala

2435

2440

2445

2450

gca aaa aaa cgg aag cgt gaa gag agt gtg gag cag aaa cgt agc aag 7869

Ala Lys Lys Arg Lys Arg Glu Glu Ser Val Glu Gln Lys Arg Ser Lys

2455

2460

2465

cag aat gcc act aag ctg tca gct ctg ctc ttc aag cac aaa gag cag 7917

Gln Asn Ala Thr Lys Leu Ser Ala Leu Leu Phe Lys His Lys Glu Gln

2470

2475

2480

ctc aga gcc gag atc ctg aag aag aga gca ctc ctg gac aag gat ctg 7965

Leu Arg Ala Glu Ile Leu Lys Lys Arg Ala Leu Leu Asp Lys Asp Leu

2485

2490

2495

caa att gaa gtg cag gaa gag ctg aag aga gac ctg aaa att aag aaa 8013

Gln Ile Glu Val Gln Glu Glu Leu Lys Arg Asp Leu Lys Ile Lys Lys
2500 2505 2510

gaa aaa gac ctg atg cag ttg gct cag gcc aca gca gta gct gca ccc 8061
Glu Lys Asp Leu Met Gln Leu Ala Gln Ala Thr Ala Val Ala Ala Pro
2515 2520 2525 2530

tgc ccc cca gtg aca cca gtt ctt cca gcc cct cca gcc cct cca cct 8109
Cys Pro Pro Val Thr Pro Val Leu Pro Ala Pro Pro Ala Pro Pro Pro
2535 2540 2545

tca cct ccc cct cca cct ggt gtg caa cac aca ggc ctt ctg tcc acg 8157
Ser Pro Pro Pro Pro Pro Gly Val Gln His Thr Gly Leu Leu Ser Thr
2550 2555 2560

ccc acc tta cct gtt gct tcc cag aag agg aag cgg gaa gag gaa aaa 8205
Pro Thr Leu Pro Val Ala Ser Gln Lys Arg Lys Arg Glu Glu Glu Lys
2565 2570 2575

gac tcc agc tca aag tcc aag aaa aag aaa atg atc tct act acc tca 8253
Asp Ser Ser Ser Lys Ser Lys Lys Lys Lys Met Ile Ser Thr Thr Ser
2580 2585 2590

aag gaa act aag aag gac aca aag ctt tac tgt atc tgt aaa acg cct 8301
Lys Glu Thr Lys Lys Asp Thr Lys Leu Tyr Cys Ile Cys Lys Thr Pro
2595 2600 2605 2610

tat gat gaa tct aaa ttt tat att ggc tgt gat cgg tgt cag aat tgg 8349
Tyr Asp Glu Ser Lys Phe Tyr Ile Gly Cys Asp Arg Cys Gln Asn Trp
2615 2620 2625

tac cat ggg cgc tgc gtt ggc atc ttg caa agt gag gca gag ctc att 8397
Tyr His Gly Arg Cys Val Gly Ile Leu Gln Ser Glu Ala Glu Leu Ile
2630 2635 2640

gat gag tat gtc tgt cca cag tgc cag tca aca gag gat gcc atg aca 8445
Asp Glu Tyr Val Cys Pro Gln Cys Gln Ser Thr Glu Asp Ala Met Thr
2645 2650 2655

gtg ctc acg cca cta aca gag aag gat tat gag ggg ttg aag agg gtg 8493
Val Leu Thr Pro Leu Thr Glu Lys Asp Tyr Glu Gly Leu Lys Arg Val
2660 2665 2670

ctc cgt tcc tta cag gcc cat aag atg gcc tgg cct ttc ctt gaa cca 8541
Leu Arg Ser Leu Gln Ala His Lys Met Ala Trp Pro Phe Leu Glu Pro
2675 2680 2685 2690

gta gac cct aat gat gca cca gat tat tat ggt gtt att aag gaa cct 8589
Val Asp Pro Asn Asp Ala Pro Asp Tyr Tyr Gly Val Ile Lys Glu Pro
2695 2700 2705

atg gac ctt gcc acc atg gaa gaa aga gta caa aga cga tat tat gaa 8637

Met Asp Leu Ala Thr Met Glu Glu Arg Val Gln Arg Arg Tyr Tyr Glu

2710

2715

2720

aag ctg acg gaa ttt gtg gca gat atg acc aaa att ttt gat aac tgt 8685

Lys Leu Thr Glu Phe Val Ala Asp Met Thr Lys Ile Phe Asp Asn Cys

2725

2730

2735

cgt tac tac aat cca agt gac tcc cca ttt tac cag tgt gca gaa gtt 8733

Arg Tyr Tyr Asn Pro Ser Asp Ser Pro Phe Tyr Gln Cys Ala Glu Val

2740

2745

2750

ctc gaa tca ttc ttt gta cag aaa ttg aaa ggc ttc aaa gct agc agg 8781

Leu Glu Ser Phe Phe Val Gln Lys Leu Lys Gly Phe Lys Ala Ser Arg

2755

2760

2765

2770

tct cat aac aac aaa ctg cag tct aca gct tct taaagttcag cgtgttaacc 8834

Ser His Asn Asn Lys Leu Gln Ser Thr Ala Ser

2775

2780

taacataaaa cacagcaaga atctggttgt ctgaactatt ttaaattaag gagccagatg 8894

tttttagtca ggctatcctg acaagacttg acctaaactt cgtttttatt ggtcataaca 8954

gtccaattat attcttggcc aattttgtcc aacggacaag aaaaaagcaa agtcaacgac 9014

accattatct tgtcaagatc agatggtttt actattgttg cagaagcgag aaaactttgt 9074

ttattgaaaa aaaaagaaaa agaaagcaag aaaaaaagat actatggggt caagtgtaac 9134

tccatgaaaa tgccacgtct gctcttcagt gaagaagctg gtttagagtc tcacagaaaa 9194

ctttigactg tatttattha ttgttgcaaa aaagacgctt ttttattgct gccctcattt 9254

gtcagctaag tattttttct tataaaatcc agccccggtt acatataatc atctgtatct 9314

tatcatgatt cctgtaggta aaagtacaag acgacctcta gatgtctttt ctttctatga 9374

aaggagctgc tatgtacaca tgtgcacaca cacacaactg ggaatcaaca atgagtttat 9434

tgttcatggt agattaaaat taagcttgca taaaggttgg gctaagtgtt ccttgggcta 9494

cagactctgt tgccttgaat ataacagtac aatttgtcaa ttactctgca ccaggctaaa 9554

gtgagtaaaa tctatttgaa ggtatcttgt ttgtaaacat ttgtcagatt ctaatttttt 9614

tcttttgtat taaaattcaa ctatggatgt atatgaaaca aaataaatgg agataatttt 9674

tctcccacaa aaaaaaaaaa aaaaaa 9700

<210> 10

<211> 2781

<212> PRT

<213> Homo sapiens

<400> 10

Met Val Ser Glu Glu Glu Glu Glu Glu Asp Gly Asp Ala Glu Glu Thr

1

5

10

15

Gln Asp Ser Glu Asp Asp Glu Glu Asp Glu Met Glu Glu Asp Asp Asp

20

25

30

Asp Ser Asp Tyr Pro Glu Glu Met Glu Asp Asp Asp Asp Ala Ser

35

40

45

Tyr Cys Thr Glu Ser Ser Phe Arg Ser His Ser Thr Tyr Ser Ser Thr

50

55

60

Pro Gly Arg Arg Lys Pro Arg Val His Arg Pro Arg Ser Pro Ile Leu

65

70

75

80

Glu Glu Lys Asp Ile Pro Pro Leu Glu Phe Pro Lys Ser Ser Glu Asp

85

90

95

Leu Met Val Pro Asn Glu His Ile Met Asn Val Ile Ala Ile Tyr Glu

100

105

110

Val Leu Arg Asn Phe Gly Thr Val Leu Arg Leu Ser Pro Phe Arg Phe

115	120	125		
Glu Asp Phe Cys Ala Ala Leu Val Ser Gln Glu Gln Cys Thr Leu Met				
130	135	140		
Ala Glu Met His Val Val Leu Leu Lys Ala Val Leu Arg Glu Glu Asp				
145	150	155	160	
Thr Ser Asn Thr Thr Phe Gly Pro Ala Asp Leu Lys Asp Ser Val Asn				
	165	170	175	
Ser Thr Leu Tyr Phe Ile Asp Gly Met Thr Trp Pro Glu Val Leu Arg				
	180	185	190	
Val Tyr Cys Glu Ser Asp Lys Glu Tyr His His Val Leu Pro Tyr Gln				
	195	200	205	
Glu Ala Glu Asp Tyr Pro Tyr Gly Pro Val Glu Asn Lys Ile Lys Val				
	210	215	220	
Leu Gln Phe Leu Val Asp Gln Phe Leu Thr Thr Asn Ile Ala Arg Glu				
	225	230	235	240
Glu Leu Met Ser Glu Gly Val Ile Gln Tyr Asp Asp His Cys Arg Val				
	245	250	255	

Cys His Lys Leu Gly Asp Leu Leu Cys Cys Glu Thr Cys Ser Ala Val

260

265

270

Tyr His Leu Glu Cys Val Lys Pro Pro Leu Glu Glu Val Pro Glu Asp

275

280

285

Glu Trp Gln Cys Glu Val Cys Val Ala His Lys Val Pro Gly Val Thr

290

295

300

Asp Cys Val Ala Glu Ile Gln Lys Asn Lys Pro Tyr Ile Arg His Glu

305

310

315

320

Pro Ile Gly Tyr Asp Arg Ser Arg Arg Lys Tyr Trp Phe Leu Asn Arg

325

330

335

Arg Leu Ile Ile Glu Glu Asp Thr Glu Asn Glu Asn Glu Lys Lys Ile

340

345

350

Trp Tyr Tyr Ser Thr Lys Val Gln Leu Ala Glu Leu Ile Asp Cys Leu

355

360

365

Asp Lys Asp Tyr Trp Glu Ala Glu Leu Cys Lys Ile Leu Glu Glu Met

370

375

380

Arg Glu Glu Ile His Arg His Met Asp Ile Thr Glu Asp Leu Thr Asn

385

390

395

400

Lys Ala Arg Gly Ser Asn Lys Ser Phe Leu Ala Ala Ala Asn Glu Glu

405

410

415

Ile Leu Glu Ser Ile Arg Ala Lys Lys Gly Asp Ile Asp Asn Val Lys

420

425

430

Ser Pro Glu Glu Thr Glu Lys Asp Lys Asn Glu Thr Glu Asn Asp Ser

435

440

445

Lys Asp Ala Glu Lys Asn Arg Glu Glu Phe Glu Asp Gln Ser Leu Glu

450

455

460

Lys Asp Ser Asp Asp Lys Thr Pro Asp Asp Asp Pro Glu Gln Gly Lys

465

470

475

480

Ser Glu Val Gly Asp Phe Lys Ser Glu Lys Ser Asn Gly Glu Leu Ser

485

490

495

Glu Ser Pro Gly Ala Gly Lys Gly Ala Ser Gly Ser Thr Arg Ile Ile

500

505

510

Thr Arg Leu Arg Asn Pro Asp Ser Lys Leu Ser Gln Leu Lys Ser Gln

515

520

525

Gln Val Ala Ala Ala Ala His Glu Ala Asn Lys Leu Phe Lys Glu Gly

530	535	540	
Lys Glu Val Leu Val Val Asn Ser Gln Gly Glu Ile Ser Arg Leu Ser			
545	550	555	560
Thr Lys Lys Glu Val Ile Met Lys Gly Asn Ile Asn Asn Tyr Phe Lys			
	565	570	575
Leu Gly Gln Glu Gly Lys Tyr Arg Val Tyr His Asn Gln Tyr Ser Thr			
	580	585	590
Asn Ser Phe Ala Leu Asn Lys His Gln His Arg Glu Asp His Asp Lys			
	595	600	605
Arg Arg His Leu Ala His Lys Phe Cys Leu Thr Pro Ala Gly Glu Phe			
610	615	620	
Lys Trp Asn Gly Ser Val His Gly Ser Lys Val Leu Thr Ile Ser Thr			
625	630	635	640
Leu Arg Leu Thr Ile Thr Gln Leu Glu Asn Asn Ile Pro Ser Ser Phe			
	645	650	655
Leu His Pro Asn Trp Ala Ser His Arg Ala Asn Trp Ile Lys Ala Val			
	660	665	670

Gln Met Cys Ser Lys Pro Arg Glu Phe Ala Leu Ala Leu Ala Ile Leu

675

680

685

Glu Cys Ala Val Lys Pro Val Val Met Leu Pro Ile Trp Arg Glu Phe

690

695

700

Leu Gly His Thr Arg Leu His Arg Met Thr Ser Ile Glu Arg Glu Glu

705

710

715

720

Lys Glu Lys Val Lys Lys Lys Glu Lys Lys Gln Glu Glu Glu Thr

725

730

735

Met Gln Gln Ala Thr Trp Val Lys Tyr Thr Phe Pro Val Lys His Gln

740

745

750

Val Trp Lys Gln Lys Gly Glu Glu Tyr Arg Val Thr Gly Tyr Gly Gly

755

760

765

Trp Ser Trp Ile Ser Lys Thr His Val Tyr Arg Phe Val Pro Lys Leu

770

775

780

Pro Gly Asn Thr Asn Val Asn Tyr Arg Lys Ser Leu Glu Gly Thr Lys

785

790

795

800

Asn Asn Met Asp Glu Asn Met Asp Glu Ser Asp Lys Arg Lys Cys Ser

805

810

815

Arg Ser Pro Lys Lys Ile Lys Ile Glu Pro Asp Ser Glu Lys Asp Glu

820

825

830

Val Lys Gly Ser Asp Ala Ala Lys Gly Ala Asp Gln Asn Glu Met Asp

835

840

845

Ile Ser Lys Ile Thr Glu Lys Lys Asp Gln Asp Val Lys Glu Leu Leu

850

855

860

Asp Ser Asp Ser Asp Lys Pro Cys Lys Glu Glu Pro Met Glu Val Asp

865

870

875

880

Asp Asp Met Lys Thr Glu Ser His Val Asn Cys Gln Glu Ser Ser Gln

885

890

895

Val Asp Val Val Asn Val Ser Glu Gly Phe His Leu Arg Thr Ser Tyr

900

905

910

Lys Lys Lys Thr Lys Ser Ser Lys Leu Asp Gly Leu Leu Glu Arg Arg

915

920

925

Ile Lys Gln Phe Thr Leu Glu Glu Lys Gln Arg Leu Glu Lys Ile Lys

930

935

940

Leu Glu Gly Gly Ile Lys Gly Ile Gly Lys Thr Ser Thr Asn Ser Ser

945	950	955	960
Lys Asn Leu Ser Glu Ser Pro Val Ile Thr Lys Ala Lys Glu Gly Cys			
	965	970	975
Gln Ser Asp Ser Met Arg Gln Glu Gln Ser Pro Asn Ala Asn Asn Asp			
	980	985	990
Gln Pro Glu Asp Leu Ile Gln Gly Cys Ser Gln Ser Asp Ser Ser Val			
	995	1000	1005
Leu Arg Met Ser Asp Pro Ser His Thr Thr Asn Lys Leu Tyr Pro Lys			
	1010	1015	1020
Asp Arg Val Leu Asp Asp Val Ser Ile Arg Ser Pro Glu Thr Lys Cys			
025	1030	1035	1040
Pro Lys Gln Asn Ser Ile Glu Asn Asp Ile Glu Glu Lys Val Ser Asp			
	1045	1050	1055
Leu Ala Ser Arg Gly Gln Glu Pro Thr Lys Ser Lys Thr Lys Gly Asn			
	1060	1065	1070
Asp Phe Phe Ile Asp Asp Ser Lys Leu Ala Ser Ala Asp Asp Ile Gly			
	1075	1080	1085

Thr Leu Ile Cys Lys Asn Lys Lys Pro Leu Ile Gln Glu Glu Ser Asp

1090

1095

1100

Thr Ile Val Ser Ser Ser Lys Ser Ala Leu His Ser Ser Val Pro Lys

105

1110

1115

1120

Ser Thr Asn Asp Arg Asp Ala Thr Pro Leu Ser Arg Ala Met Asp Phe

1125

1130

1135

Glu Gly Lys Leu Gly Cys Asp Ser Glu Ser Asn Ser Thr Leu Glu Asn

1140

1145

1150

Ser Ser Asp Thr Val Ser Ile Gln Asp Ser Ser Glu Glu Asp Met Ile

1155

1160

1165

Val Gln Asn Ser Asn Glu Ser Ile Ser Glu Gln Phe Arg Thr Arg Glu

1170

1175

1180

Gln Asp Val Glu Val Leu Glu Pro Leu Lys Cys Glu Leu Val Ser Gly

185

1190

1195

1200

Glu Ser Thr Gly Asn Cys Glu Asp Arg Leu Pro Val Lys Gly Thr Glu

1205

1210

1215

Ala Asn Gly Lys Lys Pro Ser Gln Gln Lys Lys Leu Glu Glu Arg Pro

1220

1225

1230

Val Asn Lys Cys Ser Asp Gln Ile Lys Leu Lys Asn Thr Thr Asp Lys

1235

1240

1245

Lys Asn Asn Glu Asn Arg Glu Ser Glu Lys Lys Gly Gln Arg Thr Ser

1250

1255

1260

Thr Phe Gln Ile Asn Gly Lys Asp Asn Lys Pro Lys Ile Tyr Leu Lys

265

1270

1275

1280

Gly Glu Cys Leu Lys Glu Ile Ser Glu Ser Arg Val Val Ser Gly Asn

1285

1290

1295

Val Glu Pro Lys Val Asn Asn Ile Asn Lys Ile Ile Pro Glu Asn Asp

1300

1305

1310

Ile Lys Ser Leu Thr Val Lys Glu Ser Ala Ile Arg Pro Phe Ile Asn

1315

1320

1325

Gly Asp Val Ile Met Glu Asp Phe Asn Glu Arg Asn Ser Ser Glu Thr

1330

1335

1340

Lys Ser His Leu Leu Ser Ser Ser Asp Ala Glu Gly Asn Tyr Arg Asp

345

1350

1355

1360

Ser Leu Glu Thr Leu Pro Ser Thr Lys Glu Ser Asp Ser Thr Gln Thr

1365	1370	1375
Thr Thr Pro Ser Ala Ser Cys Pro Glu Ser Asn Ser Val Asn Gln Val		
1380	1385	1390
Glu Asp Met Glu Ile Glu Thr Ser Glu Val Lys Lys Val Thr Ser Ser		
1395	1400	1405
Pro Ile Thr Ser Glu Glu Glu Ser Asn Leu Ser Asn Asp Phe Ile Asp		
1410	1415	1420
Glu Asn Gly Leu Pro Ile Asn Lys Asn Glu Asn Val Asn Gly Glu Ser		
425	1430	1435
		1440
Lys Arg Lys Thr Val Ile Thr Glu Val Thr Thr Met Thr Ser Thr Val		
1445	1450	1455
Ala Thr Glu Ser Lys Thr Val Ile Lys Val Glu Lys Gly Asp Lys Gln		
1460	1465	1470
Thr Val Val Ser Ser Thr Glu Asn Cys Ala Lys Ser Thr Val Thr Thr		
1475	1480	1485
Thr Thr Thr Thr Val Thr Lys Leu Ser Thr Pro Ser Thr Gly Gly Ser		
1490	1495	1500

Val Asp Ile Ile Ser Val Lys Glu Gln Ser Lys Thr Val Val Thr Thr
505 1510 1515 1520

Thr Val Thr Asp Ser Leu Thr Thr Thr Gly Gly Thr Leu Val Thr Ser
1525 1530 1535

Met Thr Val Ser Lys Glu Tyr Ser Thr Arg Asp Lys Val Lys Leu Met
1540 1545 1550

Lys Phe Ser Arg Pro Lys Lys Thr Arg Ser Gly Thr Ala Leu Pro Ser
1555 1560 1565

Tyr Arg Lys Phe Val Thr Lys Ser Thr Lys Lys Ser Ile Phe Val Leu
1570 1575 1580

Pro Asn Asp Asp Leu Lys Lys Leu Ala Arg Lys Gly Gly Ile Arg Glu
585 1590 1595 1600

Val Pro Tyr Phe Asn Tyr Asn Ala Lys Pro Ala Leu Asp Ile Trp Pro
1605 1610 1615

Tyr Pro Ser Pro Arg Pro Thr Phe Gly Ile Thr Trp Arg Tyr Arg Leu
1620 1625 1630

Gln Thr Val Lys Ser Leu Ala Gly Val Ser Leu Met Leu Arg Leu Leu
1635 1640 1645

Trp Ala Ser Leu Arg Trp Asp Asp Met Ala Ala Lys Val Pro Pro Gly

1650

1655

1660

Gly Gly Ser Thr Arg Thr Glu Thr Ser Glu Thr Glu Ile Thr Thr Thr

665

1670

1675

1680

Glu Ile Ile Lys Arg Arg Asp Val Gly Pro Tyr Gly Ile Arg Phe Glu

1685

1690

1695

Tyr Cys Ile Arg Lys Ile Ile Cys Pro Ile Gly Val Pro Glu Thr Pro

1700

1705

1710

Lys Glu Thr Pro Thr Pro Gln Arg Lys Gly Leu Arg Ser Ser Ala Leu

1715

1720

1725

Arg Pro Lys Arg Pro Glu Thr Pro Lys Gln Thr Gly Pro Val Ile Ile

1730

1735

1740

Glu Thr Trp Val Ala Glu Glu Glu Leu Glu Leu Trp Glu Ile Arg Ala

745

1750

1755

1760

Phe Ala Glu Arg Val Glu Lys Glu Lys Ala Gln Ala Val Glu Gln Gln

1765

1770

1775

Ala Lys Lys Arg Leu Glu Gln Gln Lys Pro Thr Val Ile Ala Thr Ser

1780	1785	1790	
Thr Thr Ser Pro Thr Ser Ser Thr Thr Ser Thr Ile Ser Pro Ala Gln			
1795	1800	1805	
Lys Val Met Val Ala Pro Ile Ser Gly Ser Val Thr Thr Gly Thr Lys			
1810	1815	1820	
Met Val Leu Thr Thr Lys Val Gly Ser Pro Ala Thr Val Thr Phe Gln			
825	1830	1835	1840
Gln Asn Lys Asn Phe His Gln Thr Phe Ala Thr Trp Val Lys Gln Gly			
1845	1850	1855	
Gln Ser Asn Ser Gly Val Val Gln Val Gln Gln Lys Val Leu Gly Ile			
1860	1865	1870	
Ile Pro Ser Ser Thr Gly Thr Ser Gln Gln Thr Phe Thr Ser Phe Gln			
1875	1880	1885	
Pro Arg Thr Ala Thr Val Thr Ile Arg Pro Asn Thr Ser Gly Ser Gly			
1890	1895	1900	
Gly Thr Thr Ser Asn Ser Gln Val Ile Thr Gly Pro Gln Ile Arg Pro			
905	1910	1915	1920

Gly Met Thr Val Ile Arg Thr Pro Leu Gln Gln Ser Thr Leu Gly Lys

1925

1930

1935

Ala Ile Ile Arg Thr Pro Val Met Val Gln Pro Gly Ala Pro Gln Gln

1940

1945

1950

Val Met Thr Gln Ile Ile Arg Gly Gln Pro Val Ser Thr Ala Val Ser

1955

1960

1965

Ala Pro Asn Thr Val Ser Ser Thr Pro Gly Gln Lys Ser Leu Thr Ser

1970

1975

1980

Ala Thr Ser Thr Ser Asn Ile Gln Ser Ser Ala Ser Gln Pro Pro Arg

985

1990

1995

2000

Pro Gln Gln Gly Gln Val Lys Leu Thr Met Ala Gln Leu Thr Gln Leu

2005

2010

2015

Thr Gln Gly His Gly Gly Asn Gln Gly Leu Thr Val Val Ile Gln Gly

2020

2025

2030

Gln Gly Gln Thr Thr Gly Gln Leu Gln Leu Ile Pro Gln Gly Val Thr

2035

2040

2045

Val Leu Pro Gly Pro Gly Gln Gln Leu Met Gln Ala Ala Met Pro Asn

2050

2055

2060

Gly Thr Val Gln Arg Phe Leu Phe Thr Pro Leu Ala Thr Thr Ala Thr
065 2070 2075 2080

Thr Ala Ser Thr Thr Thr Thr Thr Val Ser Thr Thr Ala Ala Gly Thr
 2085 2090 2095

Gly Glu Gln Arg Gln Ser Lys Leu Ser Pro Gln Met Gln Val His Gln
 2100 2105 2110

Asp Lys Thr Leu Pro Pro Ala Gln Ser Ser Ser Val Gly Pro Ala Lys
 2115 2120 2125

Ala Gln Pro Gln Thr Ala Gln Pro Ser Ala Arg Pro Gln Pro Gln Thr
 2130 2135 2140

Gln Pro Gln Ser Pro Ala Gln Pro Glu Val Gln Thr Gln Pro Glu Val
145 2150 2155 2160

Gln Thr Gln Thr Thr Val Ser Ser His Val Pro Ser Glu Ala Gln Pro
 2165 2170 2175

Thr His Ala Gln Ser Ser Lys Pro Gln Val Ala Ala Gln Ser Gln Pro
 2180 2185 2190

Gln Ser Asn Val Gln Gly Gln Ser Pro Val Arg Val Gln Ser Pro Ser

2195

2200

2205

Gln Thr Arg Ile Arg Pro Ser Thr Pro Ser Gln Leu Ser Pro Gly Gln

2210

2215

2220

Gln Ser Gln Val Gln Thr Thr Thr Ser Gln Pro Ile Pro Ile Gln Pro

2225

2230

2235

2240

His Thr Ser Leu Gln Ile Pro Ser Gln Gly Gln Pro Gln Ser Gln Pro

2245

2250

2255

Gln Val Gln Ser Ser Thr Gln Thr Leu Ser Ser Gly Gln Thr Leu Asn

2260

2265

2270

Gln Val Ser Val Ser Ser Pro Ser Arg Pro Gln Leu Gln Ile Gln Gln

2275

2280

2285

Pro Gln Pro Gln Val Ile Ala Val Pro Gln Leu Gln Gln Gln Val Gln

2290

2295

2300

Val Leu Ser Gln Ile Gln Ser Gln Val Val Ala Gln Ile Gln Ala Gln

305

2310

2315

2320

Gln Ser Gly Val Pro Gln Gln Ile Lys Leu Gln Leu Pro Ile Gln Ile

2325

2330

2335

Gln Gln Ser Ser Ala Val Gln Thr His Gln Ile Gln Asn Val Val Thr

2340

2345

2350

Val Gln Ala Ala Ser Val Gln Glu Gln Leu Gln Arg Val Gln Gln Leu

2355

2360

2365

Arg Asp Gln Gln Gln Lys Lys Lys Gln Gln Gln Ile Glu Ile Lys Arg

2370

2375

2380

Glu His Thr Leu Gln Ala Ser Asn Gln Ser Glu Ile Ile Gln Lys Gln

385

2390

2395

2400

Val Val Met Lys His Asn Ala Val Ile Glu His Leu Lys Gln Lys Lys

2405

2410

2415

Ser Met Thr Pro Ala Glu Arg Glu Glu Asn Gln Arg Met Ile Val Cys

2420

2425

2430

Asn Gln Val Met Lys Tyr Ile Leu Asp Lys Ile Asp Lys Glu Glu Lys

2435

2440

2445

Gln Ala Ala Lys Lys Arg Lys Arg Glu Glu Ser Val Glu Gln Lys Arg

2450

2455

2460

Ser Lys Gln Asn Ala Thr Lys Leu Ser Ala Leu Leu Phe Lys His Lys

465

2470

2475

2480

Glu Gln Leu Arg Ala Glu Ile Leu Lys Lys Arg Ala Leu Leu Asp Lys

2485

2490

2495

Asp Leu Gln Ile Glu Val Gln Glu Glu Leu Lys Arg Asp Leu Lys Ile

2500

2505

2510

Lys Lys Glu Lys Asp Leu Met Gln Leu Ala Gln Ala Thr Ala Val Ala

2515

2520

2525

Ala Pro Cys Pro Pro Val Thr Pro Val Leu Pro Ala Pro Pro Ala Pro

2530

2535

2540

Pro Pro Ser Pro Pro Pro Pro Pro Gly Val Gln His Thr Gly Leu Leu

545

2550

2555

2560

Ser Thr Pro Thr Leu Pro Val Ala Ser Gln Lys Arg Lys Arg Glu Glu

2565

2570

2575

Glu Lys Asp Ser Ser Ser Lys Ser Lys Lys Lys Lys Met Ile Ser Thr

2580

2585

2590

Thr Ser Lys Glu Thr Lys Lys Asp Thr Lys Leu Tyr Cys Ile Cys Lys

2595

2600

2605

Thr Pro Tyr Asp Glu Ser Lys Phe Tyr Ile Gly Cys Asp Arg Cys Gln

2610

2615

2620

Asn Trp Tyr His Gly Arg Cys Val Gly Ile Leu Gln Ser Glu Ala Glu

625

2630

2635

2640

Leu Ile Asp Glu Tyr Val Cys Pro Gln Cys Gln Ser Thr Glu Asp Ala

2645

2650

2655

Met Thr Val Leu Thr Pro Leu Thr Glu Lys Asp Tyr Glu Gly Leu Lys

2660

2665

2670

Arg Val Leu Arg Ser Leu Gln Ala His Lys Met Ala Trp Pro Phe Leu

2675

2680

2685

Glu Pro Val Asp Pro Asn Asp Ala Pro Asp Tyr Tyr Gly Val Ile Lys

2690

2695

2700

Glu Pro Met Asp Leu Ala Thr Met Glu Glu Arg Val Gln Arg Arg Tyr

705

2710

2715

2720

Tyr Glu Lys Leu Thr Glu Phe Val Ala Asp Met Thr Lys Ile Phe Asp

2725

2730

2735

Asn Cys Arg Tyr Tyr Asn Pro Ser Asp Ser Pro Phe Tyr Gln Cys Ala

2740

2745

2750

Glu Val Leu Glu Ser Phe Phe Val Gln Lys Leu Lys Gly Phe Lys Ala

2755

2760

2765

Ser Arg Ser His Asn Asn Lys Leu Gln Ser Thr Ala Ser

2770

2775

2780

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP99/02340

A. CLASSIFICATION OF SUBJECT MATTER Int.Cl ⁶ C07K14/47, C12N15/12, C12N5/10, C12P21/02, C07K16/18, C12P21/08		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) Int.Cl ⁶ C07K14/47, C12N15/12, C12N5/10, C12P21/02, C07K16/18, C12P21/08		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) REGISTRY (STN), CA (STN), MEDLINE (STN), WPI (DIALOG), BIOSIS (DIALOG), GenBank/EMBL/DBJ, SwissProt/PIR/GeneSeq		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ZHU, P. et al., "Identification and analysis of the complete cDNA sequence for the human FAC1 gene", Biochim. Biophys. Acta(1996) Vol. 1309, No. 1-2, p.5-8	1-15
A	WILSON, R. et al., "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans", Nature (1994) Vol. 368, No. 6466, p.32-38	1-15
A	BROWNELL, J.E. et al., "Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation", Cell (1996) Vol. 84, No. 6, p.843-851	1-15
A	JONES, M.H. et al., "Identification and characterization of BRDT:A testis-specific gene related to the bromodomain genes RING3 and Drosophila fish", Genomics (1997) Vol. 45, No. 3, p.529-534	1-15
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 10 August, 1999 (10. 08. 99)		Date of mailing of the international search report 17 August, 1999 (17. 08. 99)
Name and mailing address of the ISA/ Japanese Patent Office		Authorized officer
Facsimile No.		Telephone No.

国際調査報告

国際出願番号 PCT/J P 99/02340

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl.⁸ C07K 14/47, C12N 15/12, C12N 5/10, C12P 21/02, C07K 16/18, C12P 21/08

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl.⁸ C07K 14/47, C12N 15/12, C12N 5/10, C12P 21/02, C07K 16/18, C12P 21/08

最小限資料以外の資料で調査を行った分野に含まれるもの

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

REGISTRY (STN), CA (STN), MEDLINE (STN), WPI (DIALOG), BIOSIS (DIALOG),
GenBank/EMBL/DBJ, SwissProt/PIR/GeneSeq

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	ZHU, P. et al. "Identification and analysis of the complete cDNA sequence for the human FAC1 gene", Biochim. Biophys. Acta (1996) Vol. 1309, No. 1-2, p. 5-8	1-15
A	WILSON, R. et al. "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans", Nature (1994) Vol. 368, No. 6466, p. 32-38	1-15
A	BROWNELL, J. E. et al. "Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation", Cell (1996) Vol. 84, No. 6, p. 843-851	1-15

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

- 「A」 特に関連のある文献ではなく、一般的技術水準を示すもの
「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
「O」 口頭による開示、使用、展示等に言及する文献
「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

- 「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」 同一パテントファミリー文献

国際調査を完了した日

10.08.99

国際調査報告の発送日

17.08.99

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)

郵便番号 100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

高堀 栄二

4 B

9281

電話番号 03-3581-1101 内線 3448

C (続き). 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JONES, M. H. et al. "Identification and characterization of BRDT :A testis-specific gene related to the bromodomain genes RING3 and Drosophila fish", Genomics(1997) Vol. 45, No. 3, p. 529-534	1-15